

	414605	1465790_1	BE390440
	415747	155189_1	AA381209 AA381245 AA167683
	416173	1574973_1	R32782 R17318 H24152 R19676
	417742	1686282_1	R64719 Z4680 R12451
5	417974	171237_1	AA216765 T55700 H64407
	418636	177402_1	AW749655 AA226995 AW750208 AW750206
	419536	185698_1	AA603305 AA244095 AA244183
	420854	197072_1	AW269027 A684514 A1263168 AA281079
	422166	212379_1	N34524 AA305971 AW854903 AA502335 A1433430 A1203597 AW026670 AW265523 AW650787 AA317854 AW893643
10			AW835572 AW385512 A1334686 W32951 H62666 H53902 R88904 AW835732
	422696	223666_1	BE091069 BE091123 AA319569
	423633	232451_1	AW503329 AA6510 AA331571
	423841	232507_1	AW753967 AA370795 AA331630 AW862550
	423945	233566_1	AA410943 AW848953 AA334202 AA332882
15	425201	247933_1	AA352111 AW852247 AA429695
	426690	270283_1	AA382814 AA402411 AA412355
	426879	273265_1	BE069341 AW748403 AL044991 A1908240 AA393080
	430264	315008_1	AA470519 BE303010 BE302954 DE384120
	430785	329486_1	Z33201 AA486132 T72025
20	431676	335411_1	AW85464 AW971336 AA313587 AA525142
	433667	373051_1	AA743991 AA604852 AW272737
	434339	383982_1	AW754311 AA630185 AW803285
	434469	387447_1	AA634806 C18732 AA729161 AA728660
	435447	406400_1	A1872932 AA682306 BE220163 W88695 T81307 H91447
25	437152	43396_1	AL050027 BE069051
	437854	44418_1	AL118723 AL119674 A1808018 U50537
	439031	46798_1	AF075079 HA601 H48795
	439295	470321_1	BE164500 AA832198 BE164502
	444910	624951_1	A1201849 BE089007 AW946544
30	445432	63943_1	AV653771 BE089370
	446922	69865_1	BE178655 A13329 F06610 BE175602 AV661027
	447187	711623_1	R36073 A196546 R36167
	446420	76273_1	DE623004 AA330680 BE263627 BE246433
	448516	788241_1	AW898955 AW898588 AW898590 AW898663 AW898652 A1525093
35	450522	837284_1	A1988839 A1932960 A1908259
	450736	844852_1	AW970060 A1732366 A1792313 AW839644
	451024	85565_1	AA442176 AA259181
	451067	85759_1	BE172186 AA068279 AA020815 AA013437
	451340	86840_1	AW936273 AW340350 AA017208
40	452542	921410_1	AW812256 AW812257 A1906423 A1906422
	452564	92227_1	AA026777 W50065 R0961 N54721
	453472	968371_1	AL037925 AL037931 AL037957
	454307	1106070_1	AW855717 AW362452 AW362443
	454359	1130674_1	N71277 AW390764
45	454545	1223779_1	AW806399 AW868451 AW866393 AW866297 AW8517869
	454693	1228132_1	AW813428 AW813444 AW813367 AW813359 AW813429 AW813424
	454714	1230493_1	AW815098 BE154943 BE164831
	455047	1250936_1	AW882530 AW852527 AW852526
	455092	1252971_1	BE152428 AW855572 AW855607
50	455097	1253130_1	AW856802 AW855794 AW855797 AW855906 AW855796 AW855808 AW855793 AW855807
	455100	1253334_1	BE160198 AW858698 T11520 AW935930 AW856073 AW851034
	455431	1289854_1	AW838484 BE001245 BE001190
	455511	1321229_1	BE144762 AW879091
	455609	137548_1	BE011668 BE011689 BE011627 BE011079 BE011699 BE011678 BE011696 BE011675 BE011622 BE011635
55	455651	1348732_1	BE064932 BE064879 BE064853 BE064857 BE064856 BE064977 BE064990 BE064860 BE064815 BE064857 BE064804
			BE064816 BE064850 BE064805 BE064796 BE064818 BE064875 BE064819 BE064810 BE064698 BE065059
			BE066978 BE066928 BE066927
			BE068115 BE068104 BE068102 BE068096 BE068103 BE068154 BE068198
60	455708	1351284_1	BE098326 BE069280 BE069362
	455732	1353374_1	BE069390 BE067255 BE072190 BE072236
	455838	1374805_1	BE146808 BE146507 BE161803
	455935	1384144_1	BE158637 BE158688
	455945	1385688_1	BE160638 BE160606 BE160703
	456207	165078_1	AA193480
65	456482	192289_1	AA485224 AA287306 AA258121
	458094	47311_1	AF086325 W72955 W73221 AA219112
	458673	679507_1	N98626 A1302701

## TABLE 19B

5 **Table 19B** shows the genomic positioning for those pkeys lacking unigene ID's and accession numbers in Table 19. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

- 10 **Key:** Unique number corresponding to an Eex probeset  
**Ref:** Sequence source. The 7 digit numbers in the column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., *Nature* (1998) 402:489-495.  
**Strand:** Indicates DNA strand from which exons were predicted.  
**Nt\_position:** Indicates nucleotide positions of predicted exons.

	Key	Ref	Strand	Nt_position
20	400555	9801191	Minus	134694-134817
	400608	9887686	Minus	96755-97558
	400610	9887671	Minus	117605-117928,124040-124147
	400925	7661921	Plus	38183-38391,43900-44086
	401045	8117619	Plus	93044-81184,91111-91345
25	401049	7232177	Plus	149157-150992
	401093	9516157	Minus	22335-23169
	401256	9796573	Minus	45482-45620
	401283	9800093	Minus	47255-47455
	401326	9212516	Minus	226246-227505
	401418	7452889	Minus	124885-125075
30	401451	6634068	Minus	119926-121272
	401458	9167886	Plus	76485-77597
	401497	7381770	Plus	92607-92813
	401508	7534110	Minus	110779-110983
	401575	7229804	Minus	76263-76364
35	401747	9789672	Minus	118596-118816,119119-119244,119609-119761,120422-120990,130161-130381,130468-130593,131097-131268,131866-131932,132451-132575,133580-134011
	401781	7249190	Minus	63215-63435,63531-63655,63740-63901,64237-64393,64955-65037,66290-66814
	401785	7249190	Minus	163776-165956,166189-166314,166408-166589,167112-167266,167387-167469,168634-168942
	401793	7253986	Minus	102945-103063
40	401967	4006829	Plus	72693-73021,76938-77049
	402077	8117414	Plus	65014-65195
	402109	8131678	Minus	171722-171859,173197-173303
	402184	8576001	Minus	112844-112986,113505-113636
	402376	9625329	Minus	21753-22385
45	402421	8796341	Minus	49609-49662,49758-49811,86293-86346,89776-89829,90048-90101,102817-102924
	402578	9884928	Plus	66350-66496
	402608	9809429	Minus	81747-82094
	402696	7328818	Minus	23600-23731
	402765	9367767	Plus	109568-109726
50	402800	6010175	Plus	43821-44049,46181-46273
	402820	8456893	Minus	82274-82443
	402892	8065944	Minus	194384-194645
	403133	7331427	Plus	38314-38334
	403356	8569930	Plus	92339-93036
55	403388	9438331	Plus	112733-113001,114599-114735
	403426	9719529	Minus	157156-158183
	403585	8101208	Minus	131266-131769
	403593	6862650	Minus	62554-62712,69449-69802
	403637	8671936	Minus	142647-142771,145531-145762
60	403639	8671948	Plus	113294-113326,115186-115287,119649-119788
	403677	7331517	Minus	55008-55083,62860-63051
	403775	7770580	Minus	102247-102326,103095-103148
	403943	7711894	Plus	100742-100904,101322-101503

	404091	7684554	Minus	82121-83229
	404097	7770701	Plus	55512-55781
	404142	9856892	Minus	80316-80459
5	404253	9367202	Minus	55675-55055
	404274	985189	Plus	104127-104318
	404285	2326514	Plus	32282-32416
	404360	9859450	Minus	122873-122966,151324-151469,153093-153253
	404440	7528051	Plus	80430-81581
	404443	7579073	Minus	87198-87441
10	404552	7243881	Plus	19854-20010
	404561	9795980	Minus	69039-70100
	404580	6539738	Minus	240588-241589
	404721	9856648	Minus	173763-174294
	404825	8572184	Plus	47726-48046
15	404983	4432779	Minus	51178-51374,52000-52173
	405037	7543748	Minus	127374-127578
	405041	7547185	Plus	121230-121714
	405095	8072599	Plus	136877-139066
	405153	9865585	Minus	175317-176500
20	405186	7230083	Minus	135716-135851
	405232	7249042	Plus	126904-126063
	405248	7259728	Plus	637-777
	405336	8094835	Plus	33267-33563
	405384	8524123	Minus	31900-32373
25	405480	7684589	Minus	52223-52389
	405494	8060952	Minus	70284-70518
	405547	1054740	Plus	124361-124520,124914-125050
	405609	5757553	Minus	42814-43010,43583-43783,44863-45033,46429-46554,47815-48018,49961-50153,51624-51727,51823-51959,52702-52918,55469-55601,57111-57307,58189-58296,60215-60332,61482-61727
30	405638	6289229	Plus	199260-199372,199826-199929
	405654	4895155	Minus	53624-53759
	405718	8793467	Plus	113080-113285
	405822	8273498	Minus	154680-154874,155203-155379
35	405848	7851809	Minus	28135-28244
	405873	6758747	Minus	32129-32764
	405906	7705124	Minus	10835-11059
	405917	7712162	Minus	106829-107213
	405925	6758795	Plus	129935-130282
	405963	7960374	Minus	55101-55574
40	406060	9117732	Plus	68880-69374
	406151	7144605	Minus	94087-94285
	406153	9929734	Minus	12902-13069
	406182	5923850	Minus	28256-28935
45	406271	7534217	Plus	38179-38992
	406291	5686274	Plus	9562-9867
	406348	9255985	Minus	71754-71944
	406414	9259407	Plus	49593-49850
	406446	9454608	Minus	116424-116827,118721-118859,121187-121364
	406504	7711360	Minus	107068-107277
50	406554	7711566	Plus	106956-107121

**TABLE 20: 544 GENES UP-REGULATED IN BREAST CANCER COMPARED TO NORMAL ADULT TISSUES THAT ARE LIKELY TO ENCODE EXTRACELLULAR OR CELL-SURFACE PROTEINS**

Table 20 shows 544 genes up-regulated in breast cancer compared to normal adult tissues that are likely to encode extracellular or cell-surface proteins. These were selected as for Table 19, except that the ratio was greater than or equal to 3.0, the "average" normal adult tissue level was set to the 85<sup>th</sup> percentile value amongst 144 non-malignant tissues, and the 96<sup>th</sup> percentile value amongst the 73 breast cancer specimens was greater than or equal 100 units, and the predicted protein contained a structural domain that is indicative of extracellular localization (e.g. ig, fn3, egf, 7tm domains, signal sequences, transmembrane domains). The predicted protein domains are noted.

Pkey: Unique Eos probest identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 UnigenelD: Unigene number  
 Pred.ProD.Domains: Predicted Protein Domains  
 Unigene Title: Unigene gene title  
 R1: Ratio of 99<sup>th</sup> percentile of tumor to 85<sup>th</sup> percentile of normal body tissue

25	Pkey	ExAccn	UnigenelD	Pred.ProD.Domains	UnigeneTitle	R1
	408591	AF016224	Hs.46452	SS,Uteroglobulin,SS,Uteroglobulin	mammaglobin 1	168.6
	400291	AA401369	Hs.190721	TM	ESTs	73.2
	449746	AI068584	Hs.176588	SS,p450	ESTs, Weakly similar to CP47_HUMAN CYTOC	65.7
30	407277	AW170035	Hs.326736	TM	Homo sapiens breast cancer antigen NY-BR	57.6
	400292	AA250737	Hs.72472	death,ZUS,TM,Activin_recep,kinase,	BMP-R1B	56.9
	424735	U31875	Hs.272499	SS,TM	short-chain alcohol dehydrogenase family	33.8
	426878	BE068341		TM	gbtCQ3-ETC381-270100-073-c08 BT0381 Homo	60.3
	428849	NM_000230Hs.194235		SS,Lepin,SS,Lepin,	lepin (murine obesity homolog)	40.8
35	407178	AA195651	Hs.104106	SS,Dihydroorotase,	ESTs	39.3
	408000	U11690	Hs.620	Placitin_repeat,SH3,spectrin,SS,Placitin_r	bulbous pemphigoid antigen 1 (230/240kD)	37.3
	427585	D31152	Hs.179729	SS,C1q,Collagen,SS,C1q,Collagen,	collagen, type X, alpha 1 (Schmid metaph	35.2
	429441	AJ224172	Hs.204096	SS,Uterogloblin,	lipophilin B (uterogloblin family member)	30.0
	450375	AA009847	Hs.8850	SS,TM,diclinogrin,Pep_M12B_proprep,Repro	a disintegrin and metalloproteinase doma	25.7
40	420931	AF044197	Hs.100431	SS,IL8,SS	small inducible cytokine B subfamily (Cy	25.2
	422109	S73265	Hs.1473	SS,Bombesin,SS	gastrin-releasing peptide	24.8
	445730	AI264342	Hs.170042	SS,TM,Cation_efflux	ESTs	24.1
	451110	AI95040	Hs.265396	SS	ESTs, Weakly similar to transformation-r	24.0
	400297	AI127076	Hs.334473	TM	hypothetical protein DKFZp964O1278	23.8
45	420613	X51501	Hs.98949	SS,SS	prolactin-induced protein	22.8
	452744	AI257852	Hs.30504	SS,TM,GNS1_SUR4,cAMP_binding,Rila	Homo sapiens mRNA: cDNA DKFZs434E082 (tr	22.6
	424634	NM_003613Hs.151407		ig_lbp_1,SS,AAA	cartilage intermediate layer protein, nu	21.7
	420757	X78592	Hs.59915	hormone_rec,Androgen_recep,zf-C4,	androgen receptor (dihydrotestosterone r	21.7
	424389	AI056587	Hs.2833	SS	aldohyde dehydrogenase 9 family, member	20.2
50	447350	AI375572	Hs.172634	kinase,	ESTs	19.3
	456207	AA193450		SS,TM,p450,p450	gbtzz40e07.r1 Soares_NhiMPc_S1 Homo sapi	18.3
	431448	AL137517	Hs.334473	TM	hypothetical protein DKFZp964O1278	18.2
	427217	AA390272	Hs.144341	SS	ESTs	18.2
	456939	X52509	Hs.161640	SS,TM,aminotran_1_2,Cschrtn_C_term,ca	tyrosine aminotransferase	18.1
55	435496	AW040171	Hs.265396	SS	ESTs, Weakly similar to transformation-r	17.9
	402578			SS,p450,SS,TM,p450	C1001134:pl[2117372]p[165981] fatty ac	17.8
	453160	AI263307	Hs.236884	SS	H2B histone family, member L	17.8
	422505	AL120862	Hs.124165	SS	programmed cell death 8 (PCDD8)	17.7
	444342	NM_014398Hs.10887		Lamp,SS,TM,Lamp,	similar to lysosome-associated membrane	17.5

	449765	N82233	Hs.206832	SS	ESTs, Moderately similar to ALUB_HUMAN A	17.3
	428227	AA321649	Hs.2248	SS,IL8,	small inducible cytokine subfamily B (CX	17.0
	425982	D93041	Hs.155955	SS,Acetyltransferase2,	N-acetyltransferase 1 (arylamine N-acetyl	16.7
	424001	W67883	Hs.137476	kinase,	palmitoyl expressed 10	16.5
5	448595	A8014544	Hs.21672	LRRC1,LRRC2,LRRC3,serine_carpboxylase	KIA02644 gene product	16.3
	449448	D63730	Hs.57471	SS	ESTs	16.2
	418007	M13509	Hs.83189	SS,hemopexin,Peptidase_M10,SS,Peptidase	matrix metalloproteinase 1 (MMP1; Inter	15.7
	418994	AA298620	Hs.89546	SS,lectin_c,caush,EGF,SS,EGF,lectin_c,cau	selectin E (endothelial adhesion molecule	15.5
	453556	AA441839	Hs.82905	SS	hypothetical protein FL114334	15.5
10	452401	NM_007115	Hs.23932	SS,CUB,Xlink,	tumor necrosis factor, alpha-induced pro	15.0
	446591	H44188	Hs.16466	PDZ,SS	PDZ domain containing 1	14.9
	419296	AA236115	Hs.120785	SS	ESTs	14.8
	452838	U65011	Hs.30743	SS,SS	preferentially expressed antigen in mela	14.7
	422805	AA436989	Hs.121017	histone,SS,histone,histone	H2A histone family, member A	14.3
15	448390	AL035414	Hs.21068	SS	hypothetical protein	14.2
	447342	A1199288	Hs.19322	SS,Iipocollin	Homo sapiens, Similar to RIKEN cDNA 2010	14.2
	411869	W20027	Hs.23439	SS,Peptidase_M1,	ESTs	13.9
	443709	A052692	Hs.134662	SS,TM,SNF	ESTs	13.7
	459587	A4031956	Hs.130239	SS	gb2x15c04.s1 Soares_pregnant_uterus_Nih	13.7
20	442590	A1733682	Hs.130239	SS	ESTs	13.5
	400269	X07620	Hs.2258	hemopexin,Peptidase_M10,SS,Peptidase_M10	matrix metalloproteinase 10 (MMP10; str	13.5
	411589	BE339554	Hs.70937	histone,SS,histone,histone	H3 histone family, member A	13.3
	415263	AA948033	Hs.130653	SS,histone,histone,linker_histone	ESTs	13.2
	433805	AA706910	Hs.112742	SS,Ribosomal_L7Ae,	ESTs	13.1
25	407276	A1851118	Hs.326736	TM	Homo sapiens breast cancer antigen NY-BR	13.1
	443348	AW873596	Hs.182278	SS,DENN	calmodulin 2 (phosphorylase kinase, delt	13.0
	421037	A1848608	Hs.197853	SS	programmed cell death 9 (PDCD9)	12.9
	424086	A1351010	Hs.102267	SS,Lysyl oxidase	lysyl oxidase	12.8
	400255	W72838	Hs.2533	SS	aldehyde dehydrogenase 9 family, member	12.7
30	452461	N78223	Hs.108106	SS,GBA,PHD,	transcription factor	12.5
	427365	A1873274	Hs.190721	TM	ESTs	12.4
	453365	A7028544	Hs.233797	SS,TPR	ESTs	12.3
	409269	A5576933	Hs.228472	SS,TM,UPF0016,SS,TM,UPF0016	hypothetical protein FL113352	11.9
35	432596	AJ227441	Hs.278461	SS,EGF,view,SS,TM,view,	matulin 3	11.9
	408771	AW732573	Hs.47584	TM,K_tetra,ion_trans,SS,TM,K_tetra,ion_t	potassium voltage-gated channel, delayed	11.9
	432912	BE007371	Hs.200313	SS,TM,Folate_carrier	ESTs	11.9
	447033	A1357412	Hs.157601	SS	ESTs	11.8
	421155	W67879	Hs.102257	SS,Lysyl oxidase,Aldose_epim,Epimerase,	lysyl oxidase	11.8
40	424995	NM_002497	Hs.153704	kinase,SS,TM,kinase,poly(phenyl)_syn,	NIMA (never in mitosis gene a)-related k	11.7
	425396	AL049689	Hs.158358	SS	hypothetical protein similar to tansin	11.6
	438167	R28363	Hs.24286	SS,TM,7tm_1,p450,rm	ESTs	11.5
	459583	A1907673	Hs.158358	kinase,	gb2x15c04.s1 Soares_ova tumor NihHOT H	11.4
	423945	AA410943	Hs.158358	death,ZUS,TM,Activin_rec,kinase,	Homo sapiens mRNA full length insert cDN	11.4
45	439820	AL360204	Hs.283953	SS	NM_024526:Homo sapiens hypothetical prot	11.3
	402695	Hs.283953	Hs.283953	SS	KIAA1550 protein	11.2
	445293	H57645	Hs.42596	SS,Acyltransferase,	SS,hypothetical protein S24	11.1
	430217	N47863	Hs.335901	SS,RNA_pol_A,RNA_pol_A2,Ribosomal,S24	Homo sapiens c15 mRNA, partial sequence	11.1
	447184	A7026941	Hs.17516	TM,IGR	membrane-spanning 4 domains, subfamily A	11.1
	431385	BE178536	Hs.11090	SS,TM	DKFZ434G232 protein	10.9
50	423887	AL080207	Hs.134585	SS,TM,BRCT,ank,ABC_tran,ABC_tran	COB-W-like protein	10.9
	415385	R17798	Hs.7535	SS,Fork_head,	Human clone 23948 mRNA sequence	10.7
	426704	U79263	Hs.159264	SS	protein tyrosine phosphatase, receptor t	10.4
	428859	NM_007050	Hs.225852	SS,TM,Y_phosphatase,MAM,fr3,	KIAA0479 protein	10.3
	425623	AB007948	Hs.159244	SS,laminin_B,laminin_EGF,laminin_Nterm	angiostatin receptor 1	10.3
55	428368	BE440042	Hs.83326	SS,Peptidase_M10,hemopexin,SS,Peptidase	matrix metalloproteinase 3 (stromalysin	10.3
	418912	NM_000685	Hs.85472	SS,TM,7tm_1,SS,TM,7tm_1,	triple nucleotide repeat containing 9	10.3
	422026	U0736	Hs.110826	SS	ESTs	10.3
	451952	AL120173	Hs.301663	SS,kinase,	ESTs	10.2
	438199	AW016531	Hs.122147	SS,AngGap,	ESTs	10.1
60	400609	Hs.75379	Hs.75379	SS,SS,TM	C10001899:gi7506633[pr]125362 hypothe	10.0
	413472	BE242870	Hs.75379	SS	solute carrier family 1 (glut high-aff	9.9
	432374	W65815	Hs.301595	SS	Homo sapiens cDNA FL11346 fls, clone PL	9.8
	402408	Ns	Hs.301595	SS,carb_anhydride	NM_030322:Homo sapiens hypothetical pro	9.7
	445537	AJ245971	Hs.12844	SS,TM,ras	EGF-like domain, multiple 6 (EGF16)	9.7
65	451621	A1879148	Hs.26770	SS,Iipocollin,Iipocollin,	fatty acid binding protein 7, brain	9.6
	405654	Ns	Hs.26770	BT,SS	C12001521:gi7513934[pr]1731081 cca3 pr	9.5
	434988	AA18055	Hs.161160	SS	ESTs	9.5

	416220	N49776	Hs.170994	SS, TM	hypothetical protein MGC10946	9.5
	431808	M30703	Hs.270633	SS, TM, EGF, SS	amphiregulin (schwannoma-derived growth	9.5
	414142	AW363639	Hs.160042	SS, UDPGT	Homo sapiens cDNA FLJ14436 fls, clone HE	9.4
5	400238	AA032279	Hs.61655	TM	six transmembrane epithelial antigen of	9.4
	418601	AA279490	Hs.86368	SS, TM, calreticulin, SS, TM, calreticulin,	calmagin	9.4
	415539	AA733881	Hs.72472	death, ZU5, TM, Activin_recp, kinase,	BMP-R1B	9.4
	421451	AA291377	Hs.50831	TM	ESTs	9.3
	429432	AI678059	Hs.202876	SS	synaptonemal complex protein 2	9.3
	424441	AI820662	Hs.128596	SS	ESTs	9.1
10	428429	X73114	Hs.169849	SS, TM, Ig, Ig,	myosin-binding protein C, slow-type	9.1
	406867	X31126	Hs.272620	SS, Peptidase_M10, hemopexin, SS, Peptidase	SS, Peptidase_M10, hemopexin 11 (M/M-P11; stro	9.1
	448993	AW004854	Hs.228320	SS	hypothetical protein FLJ23637	9.1
	419946	AB041035	Hs.93947	Ferric_reduct, TM, Ferric_reduct,	NM_016931: Homo sapiens NADPH oxidase 4 (	9.1
	428214	H59546	Hs.126355	SS	ESTs, Moderately similar to ALU7_HUMAN A	9.0
15	427716	AT99680	Hs.25593	SS, TM, histone, Sec1, histone, sugar, tr	ESTs	9.0
	414812	X72755	Hs.77367	SS, IL8, SS, IL8	monokine induced by gamma interferon	8.8
	400295	NA		TM, ABC, Itra, ABC, membrane,	Eos Control	8.8
	422330	D30783	Hs.115263	SS, TM, EGF, SS, TM	epiregulin	8.8
20	418182	NM_004364	Hs.79069	cyclin, SS	cyclin G2	8.8
	420077	AW512260	Hs.87767	SS	ESTs	8.7
	452281	T93500	Hs.28792	SS, TGF-beta, TGF-beta, propeptide,	Homo sapiens cDNA FLJ11041 fls, clone PL	8.7
	434531	AA642007	Hs.116369	SS	ESTs	8.6
	408380	AF123050	Hs.44532	SS, TM, ubiquitin, 7ln, 3, ANF, receptor, sush	ubiquitin	8.6
	443786	AT32643	Hs.144161	TM	ESTs	8.6
25	411078	AT220220	Hs.182694	SS, SS	Coccolip	8.5
	445495	BE622941	Hs.38469	SS, SS, ENTH, L, WEQ, ENTH, L, WEQ, DNA	SS, SS, ENTH, L, WEQ, ENTH, L, WEQ, DNA	8.5
	433429	H69125	Hs.133525	SS, TM	ESTs	8.5
	424671	NM_004625	Hs.153595	SS, EGF, TM, recept, a, kd, recept, b, SS, TM, E	low density lipoprotein-related protein	8.4
30	428215	AW593419	Hs.155223	SS	stannocalcin 2	8.4
	409045	AA630662	Hs.50094	TM	Homo sapiens mRNA; cDNA DKFZP434O0515 (f	8.4
	435525	AB31297	Hs.122310	SS, p450, SS, p450	cytochrome P450, subfamily IVB, polypept	8.3
	409203	AF704743	Hs.687	SS, SS	inositol polyphosphate 4-phosphatase, ly	8.3
35	424902	NM_003668	Hs.153687	SS, Cys, Inot, SS	Norrie disease (pseudoglioma)	8.3
	431725	X85724	Hs.2839	death, ZU5, kinase, Activin_recp,	ESTs	8.3
	418092	R45154	Hs.106804	SS	GDNF family receptor alpha 1	8.2
	439640	AW449211	Hs.105445	SS, Zn, carb, Opep, Propep, M14, SS, Propep	M14 carbonylpeptidase B1 (tissue)	8.2
	427811	MB1057	Hs.160884	SS, cpm10	ESTs	8.2
	420807	AA280627	Hs.57646	SS, TGF-beta, TGF-beta, propeptide, SS	transforming growth factor, beta 2	8.2
40	428320	W47595	Hs.189300	SS, kazal,	ESTs	8.1
	447078	AW885727	Hs.301570	SS	hypothetical protein FLJ13782	8.1
	415785	AW419198	Hs.257924	SS	Homo sapiens cDNA FLJ14035 fls, clone HE	8.0
	410102	AW248508	Hs.279727	SS	Target Exon	8.0
	404547			TM	glycylserine, s1 NCLCGAP_P18 Homo sapiens	8.0
45	433687	AA743991		SS, JMPDHL, C, JMPDHL, N, CBS	ESTs	8.0
	421373	AA080829	Hs.167771	SS	CGI-52 protein	7.9
	422634	NM_016010	Hs.116821	SS, TM, SNF, SS, TM, SNF,	solute carrier family 6 (neurotransmitter	7.9
	453310	X70697	Hs.553	SS, TM	ESTs	7.8
	435957	N39015	Hs.190388	TM, cpm10, TCP1, Sema,	ESTs	7.8
50	407771	AL138272	Hs.62713	TM, TSPN, wcp, tsp, 1, EGF, thione,	ESTs	7.8
	443646	AI085198	Hs.164226	SS, TM, cadherin, Cadherin_C_term,	ESTs	7.7
	446142	AI754693	Hs.145608	SS, rna,	ESTs	7.6
	444649	AW207523	Hs.197628	SS, TM, SDF, UPAR, LY6,	Homo sapiens mRNA; cDNA DKFZP761C1712 (f	7.6
	435147	AL133731	Hs.4774	TM	ESTs	7.6
55	439138	AI742605	Hs.193696	SS	ESTs	7.5
	429220	AV207206	Hs.136319	SS, UDPGT	hypothetical protein FLJ20706	7.5
	428804	AK000713	Hs.193736	SS, SS	transcription factor AP-2 beta (activat	7.5
	453511	AL031224	Hs.33102	SS	hypothetical protein FLJ23045	7.5
	439809	R41356	Hs.101774	SS	ubiquitin 1	7.5
	414869	AA157291	Hs.21476	SS, TM, TM	LIV-1 protein, estrogen regulated	7.5
60	416275	U41060	Hs.79136	SS, Pep, M12B, propep, Reprolysin, sp, 1,	ESTs	7.4
	452882	AW378605	Hs.8687	SS, TM	duodenal cytochrome b	7.4
	452926	AI742170	Hs.31297	SS, TM, disintegrin, Pep, M12B, propep, Repro	ESTs	7.3
	453331	AI240685	Hs.8695	SS, NA, myb, DNA-binding	v-myc avian myeloblastosis viral oncogen	7.3
	420902	U22376	Hs.1334	SS	ESTs	7.2
65	450603	R43646	Hs.12422	SS, EGF, sp, 3, SS, E2F, TDP,	cartilage oligomeric matrix protein (COM	7.2
	422867	L32137	Hs.1584	SS, aldoen, SS, aldoen,	aldehyde dehydrogenase 3 family, member	7.2
	418004	U37519	Hs.87539			

5	426461	AI038165	Hs.168946	SS,GATA,	GATA-binding protein 3 (T-cell receptor	7.1
	450701	F05960	Hs.288467	SS,LRR	Homo sapiens cDNA FLJ12280 fls, clone MA	7.1
	419519	AI188719	Hs.176370	SS	ESTs	7.1
	410555	U92642	Hs.64311	TM,disintegrin,Reprolysin,	a disintegrin and metalloproteinase domain	7.1
	433138	AB026498	Hs.59729	SS,Ig,Sema,SS,Sema,afhand	semaphorin sem2	7.0
10	411558	AA102670	Hs.70725	SS,TM,SS,TM	gamma-aminobutyric acid (GABA) A receptor	7.0
	408079	W87707	Hs.82065	TM,fn3,	interleukin 6 signal transducer (gp130),	7.0
	417275	X63578	Hs.295449	SS,afhand,SS,afhand,ms	parvalbumin	7.0
	432731	R31178	Hs.267820	SS,fn3,fn1,fn2,fn2,fn1	fibronectin 1	6.9
	442818	AK001741	Hs.8738	WD40,SS	hypothalamic protein FLJ10879	6.9
15	407386	AF026942	Hs.177936	TM,IBR	gbit-homo sapiens c1g33 mRNA, partial sequ	6.8
	427427	AF077345	Hs.177936	SS,leolin_c,SS	ESTs	6.8
	410785	AW030341	SS	SS	gbit-L2-LM0079-090300-050-D03 LM0079 Homo	6.7
	401045			ICE_p20,SS,ICE_p10,ICE_p20,ICE_p10,ICE_p		6.7
		C11001883	g16753278ref(NF_039388.1) c			6.7
20	416586	AI123555	Hs.81796	SS,Reprolysin,lap_1,	ESTs	6.7
	442062	R41823	Hs.7413	TM,EPH_fod,phkinase,SAM,fn3,	ESTs; calsynenin-2	6.7
	442861	AA243637	Hs.57787	SS	ESTs	6.6
	418935	AI654599	Hs.161712	TM,Adikin_recp,phkinase,death,ZU5,	ESTs	6.6
	422060	T02093	Hs.325623	SS,TM,CD38	ESTs, Moderately similar to ALU5_HUMAN A	6.6
25	444381	BE387335	Hs.283713	SS,milo_carr	ESTs, Weakly similar to 664054 hypothet	6.6
	404081	NA		Target Exon		6.6
	417866	AW087903	Hs.82772	SS,Collagen,COLFI,TSPN,SS,TSPN	collagen, type XI, alpha 1	6.6
	428819	AI135623	Hs.183914	SS,SS	KIA05075 gene product	6.6
	410275	U85658	Hs.61795	SS,Ribosomal_S4e	transcription factor AP-2 gamma (activat	6.4
30	426236	AW067800	Hs.156223	SS	serine (or cysteine) proteinase inhibito	6.2
	419689	NM_005025Hs.78589		SS,sepin,	CD83 antigen (activated B lymphocytes, I	6.2
	416319	AI815801	Hs.79187	SS,TM,Ig,SS,TM	RAB6 interacting, kinesin-like (rakines	6.2
	412140	AA219891	Hs.73625	SS,kinesin,	ESTs	6.2
	442942	AW167087	Hs.131562	SS,Ig,Sema,phkinase,	prolactin receptor	6.1
35	448165	A4026880	Hs.25252	SS,TM,fn3,	KMFZP434G032 protein	6.1
	443162	T49851	Hs.9029	filament,SS,filament,filament	ESTs	6.1
	406802	W26713	Hs.258972	SS,TM,DAGKa,DAGKc,	cell division cycle 2, G1 to S and G2 to	6.1
	428478	Y00272	Hs.184572	SS,phkinase,phkinase	FUR2 receptor tyrosine kinase (c-erb-b2,	6.1
	400300	X03363		SS,TM,phkinase,Recep_t_domain,SH2,PH,F	ESTs	6.1
40	433404	T32982	Hs.102720	SS	glycogenin 2	6.1
	410078	U94362	Hs.55859	Glyco_transf_8,SS	Target Exon	6.1
	401781			SS,filament,Prbicosyltran,filament,Armad	adenylylate kinase 5	6.1
	447359	NM_012093Hs.18268		SS,adenylylatekinase,	Target Exon	6.1
	402230	NA		SS,TM,p50,	H2B histone family, member Q	6.1
45	427874	NM_003525Hs.2178		histone,SS,histone,	ESTs	6.0
	428398	AI249368	Hs.89558	SS,TM	metallothionein 1E (functional)	6.0
	458098	BE550224	Hs.74170	SS	interleukin 6 (interferon, beta 2)	6.0
	419958	X04430	Hs.93913	SS,S.I,IL6,	solute carrier family 16 (monocarboxylic	6.0
	416536	X32536	Hs.42845	SS,TM	ESTs	5.8
50	419705	AF033257	Hs.128151	SS,p-C2H2,	matrix metalloproteinase 9 (gelatinase B	5.8
	424637	J05070	Hs.151738	SS,Peptidase_M10,fn2,hemoexin,SS,TM,Peptidase	telod-like 1	5.8
	446679	AI823651	Hs.128700	SS	pertipin	5.8
	412196	NM_002668Hs.103263		SS	ESTs	5.7
	442117	AW684694	Hs.128689	SS,TM	LIV-1 protein, estrogen regulated	5.7
55	400303	AA242758	Hs.78136	SS,TM	KIA0862 protein	5.7
	419440	AB020899	Hs.90419	SS	ESTs, Weakly similar to ALU4_HUMAN III	5.7
	444858	AI199738	Hs.206275	SS	matrix metalloproteinase 13 (collagenase	5.6
	432239	X81334	Hs.2908	SS,Peptidase_M10,hemoexin,SS,Peptidase	ESTs	5.6
	440705	AA804244	Hs.153205	TM	ESTs	5.6
60	400288	NA		SS,TM,ABC_tran,ABC_membrane,SS	C18000922g17499103p17120903 hypothet	5.6
	446486	H39026	Hs.308	arrestin,SS	arrestin 3, retinal (X-arrestin)	5.5
	423201	NM_000163Hs.125180		SS,TM,fn3,SS	growth hormone receptor	5.5
	433043	W67554	Hs.125019	SS	lymphoid nuclear protein (LAF-4) mRNA	5.5
	438509	AF036332	Hs.58314	SS,TM,Synbasin	ESTs	5.4
65	425247	NM_005940Hs.165324		SS,Peptidase_M10,hemoexin,SS	matrix metalloproteinase 11 (MMP11; stro	5.4
	409757	NM_001898Hs.123114		SS,cystatin,	37 kDa leucine-rich repeat (LRR) protein	5.4
	425292	NM_005824Hs.155545		SS	prostate stem cell antigen	5.4
	448045	AI297436	Hs.20166	SS,TM	solute carrier family 19 (thiamine trans	5.3
	452681	AF153330	Hs.30246	SS	programmed cell death 9	5.2
	452243	AL355715	Hs.28555	SS	ESTs	5.2
	439310	AF061120	Hs.102793	SS,TM,UDPGT,casein_kappa		5.2





	441560	F13396	Ha.7868	kinase,	Homo sapiens clone 23736 mRNA sequence	4.3
	409064	AA062954	Ha.141883	,SS,CUB,	ESTs	4.3
	422667	H25642	Ha.133471	,SS,TM,FMO-like	ESTs	4.3
	454032	W31790	Ha.194293	,SS,TM	ESTs, Weakly similar to I54374 gene NP2	4.3
5	432663	AI894317	Ha.122686	TM	ESTs	4.3
	401747			,SS,Filament,filament	Homo sapiens keratin 17 (KRT17)	4.3
	432882	NM_013257	Ha.279696	kinase,kinase_C,	serum/glucocorticoid regulated kinase-II	4.2
	437036	AI571514	Ha.133022	,SS,TM	ESTs	4.2
	447754	AW073310	Ha.163533	kinase,	Homo sapiens cDNA FLJ14142 fls, clone MA	4.2
10	443194	AI864968	Ha.279009	,SS,TM	matrix Glc protein	4.2
	451871	AI821005	Ha.118699	,SS,GDNF,	ESTs	4.2
	457211	AW672585	Ha.32399	WH1,WH1	ESTs, Weakly similar to S51797 vasodilat	4.2
	421566	NM_003399	Ha.1395	z1-C2H2,SS	early growth response 2 (Krox-20 (Drosop	4.2
	421657	AI545227	Ha.105448	,SS,TM,kinase	ESTs, Weakly similar to B34087 hypopho	4.1
15	427899	AA825286	Ha.332933	,SS,SAA_proteins,ABC_membrane,ABC_tran,	serum amyloid A1	4.1
	444779	AI192105	Ha.147170	SS,TM	ESTs	4.1
	442295	AI827248	Ha.224398	,COLF1,wvc,Collagen,	Homo sapiens cDNA FLJ11469 fs, clone HE	4.1
	436396	AI683487	Ha.152213	,wnt,	wingless-type MMTV integration site fam1	4.1
	446036	AI150491	Ha.50756	,TM,Glyco_hydro_1	ESTs	4.1
20	422938	NM_001809	Ha.1594	,SS,TM,thiolase,	centromere protein A (T7K0)	4.1
	406922	S70284		,SS,TM,Desaturase,SS	gbsracyoyl-CoA desaturase [human, adipo	4.1
	439285	AL133916	Ha.172572	,SS,Ig,kinase,LRRNT,LRRCT,	hypothetical protein FLJ20093	4.1
	424800	AL035508	Ha.153203	HLH,SS	MycD family inhibitor	4.1
	429922	Z97630	Ha.226117	,SS,TM,linker_histone,7tm_1	H1 histone family, member:0	4.1
25	447178	AW594841	Ha.192417	,SS,TM	ESTs	4.0
	409038	T97490	Ha.50002	,SS,IL8,SS,IL8	small inducible cytokine subfamily A (Cy	4.0
	452747	BE153855	Ha.51480	,SS,HLH	Ig superfamily receptor LNIR	4.0
	420139	NM_003307	Ha.9535	,SS,TM,p450,	lipase, hormone-sensitive	4.0
	406877	AA479033	Ha.130315	SS,SS	ESTs, Weakly similar to A47562 B-cell gr	4.0
30	403329	NA		SS,SS	Target Exon	4.0
	438926	AW014675	Ha.137007	SS	ESTs	4.0
	430832	AI073913	Ha.100886	SS	ESTs, Weakly similar to JE0350 Anterior	4.0
	432481	AW451545	Ha.151504	,SS,Collagen,COLF1,TSPN,	Homo sapiens cDNA FLJ11973 fs, clone HE	4.0
35	452410	AL133619	Ha.29383	,SS,TM,ras	Homo sapiens mRNA; cDNA DKFZ434E2321 (f	4.0
	418661	NM_001949	Ha.1189	SS,SS	E2F transcription factor 3	4.0
	431956	X63629	Ha.2877	,SS,TM,Cadherin_C_term,cadherin,SS,TM,cad	cadherin 3, type 1, P-cadherin (placenta	4.0
	425071	NM_013989	Ha.154424	,SS,T4_deiodinase,T4_deiodinase,	deiodinase, iodothyronine, type II	4.0
40	447197	R36075		,TM,SDF,	gbv88b01.s1 Soares placenta Nb2HP Homo	4.0
	428722	U78456	Ha.190787	,SS,TIMP,	tissue inhibitor of metalloproteinase 4	3.9
	428330	L22524	Ha.2295	,SS,Peptidase_M10,SS,Peptidase_M10,	hemopematrix metalloproteinase 7 (MMP7; uterin	3.9
	423242	AL399402	Ha.125763	SS	DEM6-6 protein	3.9
	446046	Z45051	Ha.22920	,SS,SS,TM	similar to S69401 (cattle) glucose induc	3.9
	414831	NI1169	Ha.77439	,SS,cAMP_binding,Rila,HLK_box	protein kinase, cAMP-dependent, regula	3.8
	413569	AW462631	Ha.133803	,SS,aldehydease	ESTs, Highly similar to AF157633 1 noncl	3.8
45	406875	NM_015434	Ha.48604	SS	DKFZP434B168 protein	3.8
	418629	BE247550	Ha.88669	,SS,SH2,PH,SS,TM,PH,SH2,Furin-like,kinas	growth factor receptor-bound protein 7	3.8
	450787	AB005190	Ha.25475	,SS,TM,MIP,SS,TM,MIP,	aquaporin 7	3.8
	414870	N72264	Ha.300670	SS	XKAI204 protein	3.8
	450325	AI935962	Ha.262689	SS	ESTs	3.8
50	407633	NM_007069	Ha.37169	TM,TM	similar to rat HREV107	3.8
	428172	AA371307	Ha.125056	,SS,DENN	ESTs	3.8
	442262	BE170851	Ha.8790	,SS,START,	deleted in liver cancer 1	3.8
	427961	AW293165	Ha.143134	,SS	ESTs	3.8
55	445683	AW673605	Ha.149006	,SS,WH1,WH1	ESTs	3.8
	403943			p450,SS,p450	C8000356/g[4503225]r[NP_000765.1] cyt	3.8
	408761	AA057264	Ha.238536	,SS,TM,7tm_1,	ESTs, Weakly similar to (define not av	3.8
	423278	AW569661	Ha.239843	SS	ESTs	3.8
	420440	NM_002407	Ha.97944	,SS,SRRC,Uteroglobin	memmaglobin 2 (MGB2; memmaglobin B; lip	3.7
	445107	AI208121	Ha.147313	,SS,TM	ESTs, Weakly similar to I38022 hypopho	3.7
60	426303	AW974476	Ha.183601	,SS,RGS,RGS,RGS	regulator of G-protein signalling 16	3.7
	411957	BE160198		TM	gbtCV1-HT0413-010200-059-h03 HT0413 Homo	3.7
	427809	M26360	Ha.180678	,SS,lipase,PLAT,Sec7,PH,	lipoprotein lipase	3.7
	418203	X54942	Ha.83758	,CKS,SS,CKS,	CDC26 protein kinase 2	3.7
	430376	AW292053	Ha.12532	SS	chromosome 1 open reading frame 21	3.7
65	444190	AI878918	Ha.10526	SS	cysteine and glycine-rich protein 2	3.7
	433495	AW373784	Ha.71	,SS,Ig,MHC_I,connexin,SCAN,SS,TM	alpha-2-glycoprotein 1, zinc	3.7
	429638	AI816662	Ha.211577	,SS,TM,SS	kinectin 1 (kinectin receptor)	3.7

	454071	AI041793	Ha.42602	.TM7tm_1	ESTs	3.7
	451860	HA4491	Ha.252938	.SS.TM.EGF_idl_recept_a_idl_recept_b_EGF	ESTs, Weakly similar to ALU1_HUMAN ALU 8	3.7
	423281	AI623663	Ha.191633	.SS.AAA	ESTs	3.7
	427691	AW194426	Ha.20726	.SS.Glycos_transf_2	ESTs	3.7
5	426824	W23624	Ha.173059	SS	ESTs	3.7
	424676	Y0585	Ha.151678	Glycos_transf_2_Ricin_B_lectin,SS.Glycos	UDP-N-acetyl-alpha-D-galactosamine:poly	3.7
	418026	BE379727	Ha.83213	lipocalin,SS,lipocalin,lipocalin,feritil	fatty acid binding protein 4, adipocyte	3.7
	457495	AW301344	Ha.122808	.SS.Phiboxylran,Sulfatase	DNA replication factor	3.7
	417601	NM_014735	Ha.82292	PHD,phkinase,SS	KIAA0215 gene product	3.7
10	407999	AI126271	Ha.49433	SS	ESTs, Weakly similar to Y228_HUMAN HYPOT	3.7
	425548	AA890023	Ha.1905	.SS,TM,fn3,SS,TM,fn3	prolactin receptor	3.7
	448619	AU076843	Ha.313	.SS,TM,ehand,ion_trans	secreted phosphoprotein 1 (osteopontin,	3.7
	411213	AA676939	Ha.89285	.SS,TM,CUB,F5_F8_type_C,MAM,SS,TM,CUB,F5	neurophilin 1	3.6
	406825	Y13647	Ha.119597	.SS,TM,Desaturase,SS	stearoyl-CoA desaturase (delta-9-desatur	3.6
15	417511	ALD49176	Ha.82223	SS	chordin-like	3.6
	426769	AW207175	Ha.106771	.SS,7tm_1,SPRY	ESTs	3.6
	401737	T97307		.SS,TM,CDM1_CD39	glycyls3h05.s1 Soares fetal liver spleen	3.6
	401866			.SS,filament	Target Eion	3.6
20	451195	U10492	Ha.438	.SS,homobox,Ests,SS,homobox	mesenchyme homeo box 1	3.6
	426044	AA502490	Ha.336695	SS	ESTs	3.6
	426310	NM_000609	Ha.169266	.SS,TM,7tm_1	neuropeptide Y receptor Y1	3.6
	440029	AW089705	Ha.238711	SS	ESTs, Weakly similar to S64329 probable	3.6
	408573	AA284775	Ha.43148	.SS,TM,PMF22_Claudin	ESTs	3.6
25	431830	Y18645	Ha.271387	.SS,TM,IL8	small inducible cytokine subfamily A (Cy	3.6
	444781	NM_014400	Ha.11950	.SS,PH,lactamase_B	GPI-anchored metastasis-associated prot	3.6
	431493	AI791493	Ha.123873	.SS,p450,p450	ESTs, Weakly similar to A36038 cytochrome	3.6
	414175	AI308876	Ha.103849	.TM,homoxipin,Peptidase_M10,homoxipin,Pe	hypothetical protein DKFZp761D112	3.6
	411789	AF245005	Ha.72157	.lg.LRRCT	DKFZp554I1922 protein	3.6
30	418851	AH17828	Ha.152435	.SS,TM	ESTs	3.5
	453989	AA847849	Ha.82711	.SS,HMG_box	Homo sapiens, clone IMAGE:3351295, mRNA	3.5
	407104	SE7295	Ha.323810	.SS,TM,SH2,PH,phkinase,Recept_1_domain,Fun-arb-b2	cvtan erythroleukemia	3.5
	449051	AW961400	Ha.333625	SS	HER2 receptor tyrosine kinase (c-erb-b2,	3.5
	434398	AA121098	Ha.3838	phkinase,POLO_box,SS,phkinase,POLO_box	serum-inducible kinase	3.5
35	454042	H22570	Ha.172572	.SS,lg,phkinase,LRRTN,LRRTCT	hypothetical protein FLJ20093	3.5
	459496	AA808940	Ha.274450	.SS,TM,KRAB,SCAN,zf-C2H2,lg	EST	3.5
	414998	NM_002543	Ha.77729	.SS,TM	oxidized low density lipoprotein (lectin	3.5
	407756	AA116021	Ha.38260	.SS,UCH-1,UCH-2,SS,TM,g_u_transpopt	ubiquitin specific protease 18	3.5
	442101	AI351930	Ha.135894	SS	ESTs	3.5
	449722	BE280074	Ha.23980	cyclin,SS,TM,cyclin	cyclin B1	3.5
40	452534	AW462434	Ha.58006	SS	ESTs, Weakly similar to ALU5_HUMAN ALU 5	3.5
	421991	NM_014918	Ha.10488	SS	KIAA0990 protein	3.4
	420058	AK001423	Ha.34694	SS	Homo sapiens cDNA FLJ10051 fls, clone NT	3.4
	425776	U25125	Ha.159499	.SS,TM,7tm_2,SS,TM,7tm_2	parathyroid hormone receptor 2	3.4
45	407846	AA428202	Ha.40403	.TM,ABC_membrane,ABC_1tan,Ribosomal_S4eCbp,300S-interacting transactivator, wit		3.4
	408925	L34041	Ha.9739	.SS,TM,transport_prot,SWIB,RhoGAP,DAG_Figlypon-3-phosphoinositide dehydrogenase 1 (so		3.4
	445873	AA265070	Ha.251946	.SS,mm,PABP,phkinase,14-3-3,mm	poly(A)-binding protein, cytoplasmic 1-4	3.4
	416054	NM_002318	Ha.83354	.SS,TM,mto_carr,Lysyl_oxidase	lysyl oxidase-like 2	3.4
	414921	BE390551	Ha.77628	.SS,START,SS,START,NM1T_PNMT_TENT	steroidogenic acute regulatory protein r	3.4
50	452288	NM_003512	Ha.28777	.SS,histone,Calc_CGRP_APP,lg,MHC-1,SPRY	H2A histone family, member L	3.4
	428862	NM_003458	Ha.2316	.SS,HMG_box	SPRY (sex determining region Y)-box 9 (ca	3.4
	412520	AA442334	Ha.795	histone,SS,histone,BolA	H2A histone family, member O	3.4
	410530	M25809	Ha.64173	ATP-synt_lab,SS,7tm_1,ATP-synt_lab	ATPase, H transporting, lysosomal (vacuo	3.4
	401780			filament,SS,filament,filament	NM_005557*Homo sapiens keratin 16 (foca	3.4
55	447131	NM_004685	Ha.17466	TM	retinoic acid receptor responder (tezaro	3.4
	418334	AA319233	Ha.5521	.SS,TM,Ribosomal_L27e	ESTs	3.4
	416138	C18395	Ha.235844	.Kunitz_BPTL	tissue factor pathway inhibitor 2	3.4
	421168	AF182277	Ha.330780	.SS,p450,SS	cytochrome P450, subfamily J1B (phenobar	3.4
	431473	AA826886	Ha.321176	SS	ESTs, Weakly similar to S55824 reverse t	3.4
60	421379	Y15221	Ha.903682	.SS,IL8	small inducible cytokine subfamily B (Cy	3.4
	411984	NM_005419	Ha.72988	SH2,STAT,SS,STAT	signal transducer and activator of trans	3.4
	408101	AW968504	Ha.123073	phkinase	CDCC-related protein kinase 7	3.4
	405366			RhoGEF,PH,SS,RhoGEF,PH	NM_003371*Homo sapiens vav 2 oncogene (	3.4
	414612	BE274552	Ha.76578	SAP,SS-FG-GAP,wwa	protein inhibitor of activated STAT3	3.4
65	411393	AW797437	Ha.69771	.SS,sushi,lypsin,wwa,rm, fibrinogen_C,fn	B-factor, propenidin	3.3
	435767	H73505	Ha.118784	.SS,Peptidase_S8,PH,Peptidase_S8,PH	ESTs	3.3
	416405	D86951	Ha.79299	.SS,TM	lipoma HMGIC fusion partner-like 2	3.3
	433088	NM_005458	Ha.288215	.SS,Priboxylran	sialyltransferase	3.3

5	445462	AA378776	Hs.288649	SS,SS	hypothetical protein MGC3077	3.3
	439452	AA918317	Hs.57967	SS,SS	B-cell CLL/lymphoma 118 (cib finger pro	3.3
	452017	AF109302	Hs.27495	SS	prostate cancer associated protein 7	3.3
	409089	AK000725	Hs.50579	SS	hypothetical protein FLJ20718	3.3
	452108	A1141031	Hs.21342	SS	ESTs	3.3
10	447619	U46258	Hs.339665	SS	ESTs	3.3
	428928	AF037062	Hs.172914	SS,adh_short,TGF-beta,TGF-beta_propeptide	retinol dehydrogenase 5 (11-cis and 9-cis	3.3
	438825	BE327427	Hs.78953	SS,TM,histone,ANF_receptor,guanylate_cy	ESTs	3.3
	414575	H11257	Hs.22968	SS,pkinase,Ig	Homo sapiens clone IMAGE461939, mRNA se	3.3
	417837	AL079905	Hs.1103	SS,TGF-beta_propeptide,TGF-beta,SS	transforming growth factor, beta 1	3.3
15	422126	AW881145		SS	gb:QV0-OT0033-010400-182-a07 OT0033 Homo	3.3
	445941	A1287371	Hs.172636	SS,SS,ipoxycyclo-oxygenase,PLAT	ESTs	3.3
	428973	AI423317	Hs.164680	SS,T-box,UOPGT	ESTs	3.3
	444542	A1161293	Hs.280390	SS,SS,Peptidase_M1,EGF,Ig,Ig_c,sushi	aminopeptidase	3.3
	459681	A1647306	Hs.134981	SS	ESTs	3.3
20	425741	AF052152	Hs.158412	pkinase,	Homo sapiens clone 24629 mRNA sequence	3.3
	426501	AW043782	Hs.293616	SS	ESTs	3.3
	459503	AA502764	Hs.123499	SS	ESTs, Weekly similar to AF208655 1 BM-01	3.3
	434228	Z42047	Hs.283978	SS,TM,7tm_1	Homo sapiens PRO2751 mRNA, complete cds	3.3
	415752	BE314524	Hs.78776	TM	putative transmembrane protein	3.3
25	400419	AF084545		SS,Peptidase_M1,	Target	3.3
	439750	AL369053	Hs.57664	TM,Integrin_B,Ricin_B,Ig_c,lin,rm	Homo sapiens mRNA full length insert cDN	3.3
	423858	AL137329	Hs.134363	SS,TM	Homo sapiens mRNA; cDNA DKFZ43460550 (f	3.3
	428514	AW238861	Hs.193139	SS,START,NNMT_PNMT_TEMT,	ESTs	3.3
	428698	AA852773	Hs.334836	SS	KIAA1866 protein	3.3
30	448988	Y08763	Hs.22785	SS,TM,TM	gamma-aminobutyric acid (GABA) A recepto	3.3
	432072	N62937	Hs.269109	Sema,Ig,	ESTs	3.3
	417433	BE270286	Hs.81228	SS,TM,LRRC7,LRRC7,LRRC7,TM,LRRC7,	574 oncotel trophoblast glycoprotein	3.3
	452194	AI084413	Hs.332649	SS,TM,7tm_3,ANF_receptor,sushi	olfactory receptor, family 2, subfamily	3.2
	440451	N48373	Hs.10247	SS,Ig,	activated leucocyte cell adhesion molecu	3.2
35	420042	AW015140	Hs.161723	SS,CUB,	ESTs	3.2
	467292	A1921270	Hs.334892	SS,TM,SS,TM,G-patch	hypothetical protein FLJ14251	3.2
	421458	NM_005684	Hs.194576	SS	carbohydrate (keratan sulfate Gal-6) sul	3.2
	431104	AW070859	Hs.313503	Sema,Ig,	ESTs	3.2
	437687	BE592136	Hs.9736	SS,PC1,RasGEF,hormona_rec,zf-C4,	proteasome (prosome, macropain) 28S subu	3.2
40	419589	AW973708	Hs.201925	FGF,	Homo sapiens cDNA FLJ13346 fs, clone PL	3.2
	415447	Z97171	Hs.78454	SS,OLF,OLF,OLF,Ribosomal_L4	myocilin, trabecular meshwork inducible	3.2
	434364	BE548446	Hs.5157	SS,TM,SSF,SS,TM	Homo sapiens mRNA; cDNA DKFZp34F152 (f	3.2
	423431	AA326082		SS,p450,p450	gb:EST29171 Cerebellum II Homo sapiens c	3.2
	413278	BE583085	Hs.833	SS,TM,ubiquitin,laminin_G,laminin_EGF,k	interferon-stimulated protein, 15 kDa	3.2
45	458451	AW297181	Hs.185922	SS,Ribosomal_L14	ESTs	3.2
	440449	AA885430	Hs.201925	FGF,	Homo sapiens cDNA FLJ13346 fs, clone PL	3.2
	413753	U17760	Hs.75517	SS,laminin_EGF,laminin_V,amn,adh_short,S	laminin, beta 3 (nicotin (125KD), kalinin	3.2
	434876	AF180477	Hs.61480	SS,HLH	Ig superfamily receptor LNRR	3.2
	435575	AF213457	Hs.44234	SS,Ig,SS,TM	triggering receptor expressed on myeloid	3.2
50	415773	T21691	Hs.324725	SS,TM,Ribosomal_S24e,G-gamma	ESTs, Moderately similar to AF4582 B-cell	3.2
	446440	AV858411	Hs.42655	SS	KIAA1651 protein	3.2
	450847	NM_003155	Hs.25590	SS,homobox,	stannocalcin 1	3.2
	428075	AW513691	Hs.270149	SS,fn3,	ESTs, Weekly similar to 2109250A B cell	3.2
	452110	T47567	Hs.28005	SS,TM,Activin_recp,plnase	Homo sapiens cDNA FLJ11309 fs, clone PL	3.2
55	439963	AW247529	Hs.6793	TM,p450,Est	platelet-activating factor acetylhydrola	3.2
	462837	NA		SS	ENSPC000024132-DJ947L8.1.8 (novel) Sush	3.2
	438451	AF086270	Hs.278554	SS,Chromo_shadow,chromo,	heterochromatin-like protein 1	3.1
	406664	L34041	Hs.9739	SS,TM,transpor_prc1,SWIB,RhoGAP,DAG	Pegylcerol-3-phosphate dehydrogenase 1 (so	3.1
	417315	AI080042	Hs.339801	SS,RNA_pol_A,RNA_pol_A2,Ribosomal_S24e,ribosomal protein S24	bicycan	3.1
60	413011	AW068115	Hs.821	SS,LRRC7,LRRC7,SS,LRRC7,LRRC7,	hypothetical protein FLJ14950	3.1
	414987	AA524394	Hs.294022	SS,connexin,hormona_rec,zf-C4,connexin	ESTs, Weekly similar to T202727 hypothet	3.1
	422197	H24471	Hs.28930	SS,Gelsolin,	membrane-spanning 4-domains, subfamily A	3.1
	448030	N30714	Hs.325960	SS,TM	collagen, type VII, alpha 2	3.1
	407604	AW191982	Hs.249239	SS,TM,C1q,	much 1, transmembrane	3.1
65	419092	JO561	Hs.89903	SS,TM,SEA,	Homo sapiens, clone MGC16327, mRNA, com	3.1
	458572	AK002016	Hs.114727	SS,PKC_PK	TONDU	3.1
	443711	BE281126	Hs.9630	SS,TM,7tm_1,rm,SS	Homo sapiens cDNA FLJ10071 fs, clone HE	3.1
	452256	AK000853	Hs.28661	TM,GDI7tm_1,	Transmembrane protease, serine 3	3.1
	432201	AJ538613	Hs.295241	SS,TM,Igypsin,SS,TM,tetrol,trypsin,lref	gb:Homo sapiens mRNA for immunoglobulin	3.1
	408942	AJ245210		SS	Target Exon	3.1
	409093	NA		SS		3.1

	434406	AI031771	Hs.132588	,SS,Glyco_hydro_2	ESTs	3.1
	452994	AW662597	Hs.31305	SS,WD40,SS,WD40,	KIAA1547 protein	3.1
	445803	AI347487	Hs.132781	fn3,SS,TM,EF1BD	class I cytokine receptor	3.1
	424364	AW383226	Hs.201189	SS	ESTs, Weakly similar to G01763 atrophin-	3.1
5	410196	AI596442	Hs.59838	UBACT_repeat,SS,UBACT_repeat,THIF_family	hypothetical protein FLJ10808	3.1
	419190	T29618	Hs.89640	,TM,phosphatase,fn3,	TEK tyrosine kinase, endothelial (venous)	3.1
	433417	AA587773	Hs.8859	,SS,SRGR,	Homo sapiens, Similar to RIKEN cDNA.5830	3.1
	418624	AI734080	Hs.104211	,Sema,lg,	ESTs	3.1
	436291	BE568462	Hs.5101	,SS,abhydrolase,	protein regulator of cytokinesis 1	3.1
10	411000	N04049	Hs.201619	SS	ESTs, Weakly similar to S33883 SEB4B pro	3.1
	412519	AA192421	Hs.73980	SS,Troponin,Hemagglutinin,SS,TM,C2,Tropo	troponin T1, skeletal, slow	3.1
	450223	AA418204	Hs.241493	,SS,pro_isomerase,	natural killer-tumor recognition sequenc	3.1
	422780	AA605873	Hs.25933	,TM,Helicase,Sec1,Helicase,sugar_tr	ESTs	3.1
	424269	AW137691	Hs.199754	,SS,TM,Tm_2,GPS	ESTs	3.1
15	435854	AJ278120	Hs.4995	,SS,WD40	putative ankyrin-repeat containing prote	3.1
	447389	AW630534	Hs.76277	,SS,TM,mm,oxidoreductase,oxidoreductase	Homo sapiens, clone MGC-8381, mRNA, comp	3.1
	451631	R08866	Hs.76277	SS	ghy97802.1? Scavenger fetal liver spleen	3.0
	448105	AW591433	Hs.298241	,SS,TM,tetrol,tryptase,tetrol	Transmembrane protease, serine 3	3.0
20	438637	BE500941	Hs.126730	,TM,PH,	ESTs, Weakly similar to KIAA1214 protein	3.0
	423024	AA593731	Hs.325823	,SS,TM,CD36	ESTs, Moderately similar to ALU5_HUMAN A	3.0
	456592	R91800	Hs.162209	,SS,Ran_BP1,LIM,Ran_BP1,GRIP,TPR,LIM	ghy91002.1? Scavenger fetal liver spleen	3.0
	425920	AL049977	Hs.162209	,SS,TM,SS,TM,PMP22_Claudin,PMP22_Claudin	claudin 8	3.0
	444670	H56373	Hs.332938	,SS,TM	hypothetical protein MGC5370	3.0
	401197			art,art,	ENSP0000022623*HSPC213.	3.0
25	437755	AW204256	Hs.291857	,wnt,	ESTs	3.0
	452590	BE077084	Hs.336432	,SS,mm,zf-RanBP,phosphatase,C2,phosphatase_C,DA	ESTs	3.0
	410274	AA361807	Hs.61762	SS,SS	hypoxia-inducible protein 2	3.0
	450056	V27249	Hs.8109	SS	hypothetical protein FLJ21080	3.0
	404825			,SS,TM	Target Exon	3.0
30	456369	H70284	Hs.160152	,SS,RA	ESTs, Weakly similar to FPHU alpha-1-olap	3.0
	408196	AL034548	Hs.43627	HMG_box,phosphatase,zf-CCHC,SS,TM,HMG_box,	SRV (sex determining region Y)-box 22	3.0
	433675	AW977663	Hs.76319	,SS,ribonuc_red_sm,	ribonucleotide reductase M2 polypeptide	3.0
	418843	AI820961	Hs.193465	,death,ZUS,phosphatase,Activin_reap,	ESTs	3.0
	422095	AI868872	Hs.282804	,SS,Cu-oxidase,SS,Cu-oxidase,Cu-oxidase	hypothetical protein FLJ22704	3.0
35	415992	C05837	Hs.145807	,SS,TM	hypothetical protein FLJ13693	3.0
	424631	AA688021	Hs.179808	SS	ESTs	3.0
	409596	AW103564	Hs.727	SS,TGF-beta,TGF-beta_propeptide,SS,TGF-beta,	inhibin, beta A (activin A, activin AB a	3.0
	419657	AU077005	Hs.92208	,SS,disintegrin,Reprolysin,Pep_M12B_prope	a disintegrin and metalloproteinase doma	3.0
	450946	AA374589	Hs.127698	SS	ESTs, Moderately similar to 2109260A B c	3.0
40	447770	AB032417	Hs.18545	Fritzzled,Fz,SS,TM,Fritzzled,Fz,	fritzzled (Drosophila) homolog 4	3.0
	423628	U20325	Hs.1707	SS,SS	cocaine- and amphetamine-regulated trans	3.0
	418836	AW385224	Hs.35186	,SS,TM,Phosphodiester,	ectonucleotide pyrophosphatase/phosphodi	3.0
	442804	AW300118	Hs.131257	,SS,TM,gamma	ESTs	3.0
	432284	AA532807	Hs.105822	,SS,TM,phosphatase,	ESTs	3.0

TABLE 20A

Table 20A shows the accession numbers for those pkeys lacking unigeneID's for Table 20. For each probe set, we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Prkey: Unique Eas probe set identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey	CAT number	Accessions
410785	1221055_1	AW803341 AW803265 AW803403 AW803465 AW803402 AW803413 AW803268 AW803396 AW803334 AW803365
411667	1263334_1	BE180198 AW835898 T11520 AW835930 AW866073 AW861034
418636	177402_1	AW749855 AA225995 AW750208 AW750206
420854	197072_1	AW296927 AB64514 AIZ63168 AA281079
422128	211994_1	AW861145 AW60718 M53537 AA304575 T06067 AA331991
423431	228162_1	AA328062 AA325758 AW962162
425945	233565_1	AA410943 AW848953 AA334202 AA332882
426578	273265_1	BE069341 AW748403 ALD44891 AJ908240 AA393080
433687	373061_1	AA743991 AA604852 AW772737
447197	711623_1	R38075 A1366546 R36167
451631	678098_1	R08066 R01523 A1806815
456207	165078_1	AA193450
456592	202694_1	R91800 T87079 AA291455

TABLE 20B

Table 20B shows the genomic positioning for those pkeys lacking unigene ID's and accession numbers in Table 20. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

**Pkey:** Unique number corresponding to an Eos probe set  
**Ref:** Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1993) 402:489-495.  
**Strand:** Indicates DNA strand from which exons were predicted.  
**Nt\_position:** Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
400608	9887686	Minus	96756-97558
400603	2911732	Plus	59112-59228
401045	8117619	Plus	90044-90184,91111-91345
401093	8516137	Minus	22335-23186
401197	9719705	Plus	176341-176452
401747	9789672	Minus	118666-118816,119119-119244,119609-119761,120422-120960,130161-130381,130468-130593,131057-131258,131865-131932,132461-132575,133590-134011
401760	7249190	Minus	28397-28617,28920-29045,29135-29266,29411-29567,29705-29767,30224-30573
401781	7249190	Minus	83215-83435,83531-83656,83740-83901,84237-84393,84955-85037,86290-86814
401866	6018106	Plus	73126-73623
402230	9966312	Minus	25782-25932
402408	9796239	Minus	110326-110491
402578	9884928	Plus	65330-66496
402605	9939429	Minus	81747-82094
402837	9389121	Minus	2013-2186,9570-9758,11136-11309,19429-19677,21210-21455,23368-23562,24342-24527,29132-29320
403199	9956183	Minus	56895-59036,66618-66789
403329	8516120	Plus	96450-96599
403593	6862650	Minus	62554-62712,69449-69602
403943	7711864	Plus	100742-100804,101322-101503
404091	7684554	Minus	82121-83229
404347	9838195	Plus	74493-74829
404826	6572184	Plus	47726-48046
405366	2182280	Plus	22478-22632
406854	4895165	Minus	53624-53769

**TABLE 21: 210 GENES UP-REGULATED IN BREAST CANCER COMPARED TO NORMAL ADULT TISSUES THAT ARE LIKELY TO ENCODE EITHER ENZYMES OR PROTEINS AMENABLE TO MODULATION BY SMALL MOLECULES**

Table 21 shows 210 genes up-regulated in breast cancer compared to normal adult tissues that are likely to encode either enzymes or proteins amenable to modulation by small molecules. These were selected as for Table 19, except that the ratio was greater than or equal to 3.0, the "average" normal adult tissue level was set to the 85<sup>th</sup> percentile value amongst 144 non-malignant tissues, and the 96<sup>th</sup> percentile value amongst 73 breast cancers was greater than or equal 80 units, and the predicted protein contained a structural domain that is indicative of enzymatic function or of being modulatable by small molecules (e.g. pkinase, peptidase, phosphatase, ATPase, or ion\_transporter domains). The predicted protein domains are noted.

Pkey: Unique Eos probest identifier number  
 ExAccn: Exemplar Accession number, Genbank accession number  
 Unigenid: Unigene number  
 Unigene Title: Unigene gene title  
 R1: Ratio of 93<sup>rd</sup> percentile tumor to 85<sup>th</sup> percentile of normal body tissue

Pkey	ExAccn	Unigenid	Predicted Protein Domains	UnigeneTitle	R1
449746	AI668594	Hs.176588	SS,p450	ESTs, Weakly similar to CP4Y_HUMAN CYTOD	65.7
400292	AA250737	Hs.72472	death,ZU5,TM,Actin_recp,kinase,	BMP-R1B	55.9
424735	U31875	Hs.272499	SS,TM	short-chain alcohol dehydrogenase family	53.8
407178	AA195851	Hs.104108	SS,Dihydroorotase,	ESTs	39.3
406045	AW138559	Hs.245123	Phosphodiester3,Somatostatin_B,	ESTs	34.9
450375	AA003647	Hs.2850	SS,TM,dialinigin,Pep_M12B,procep,Repro	a disintegrin and metalloproteinase doma	25.7
429170	NM_001394	Hs.2359	DSPc,Rhodanese,	dual specificity phosphatase 4	24.9
445730	AI624342	Hs.170042	SS,TM,Cation_efflux	ESTs	24.1
424634	NM_003613	Hs.151407	lg,isp_1,SS,AAA	cartilage intermediate layer protein, nu	21.7
420757	X78592	Hs.99615	hormone_rec,Androgen_recep,zf-C4,	androgen receptor (dihydrotestosterone r	21.7
424359	AI605687	Hs.2533	SS	aldehyde dehydrogenase 9 family, member	20.3
447350	AI375672	Hs.172634	kininase,	ESTs	19.2
456207	AA193450		SS,TM,p450,p450	gbzr40e07.r1 Soares_NHMPu_S1 Homo ssp	18.3
456938	X52509	Hs.161640	SS,TM,aminotran_1_2,Cadherin_C_term,	cadhytyrosine aminotransferase	18.1
402578			SS,p450,SS,TM,p450	C1001134:g12117372p1r1j65881 fatty ac	17.8
425682	D90041	Hs.155956	SS,Acetyltransf2,	N-acetyltransferase 1 (arylamine N-acety	16.7
424001	W67883	Hs.137476	kininase,	paternally expressed 10	16.5
418007	M13529	Hs.83169	SS,hemopexin,Peptidase_M10,SS,Peptidase	matrix metalloproteinase 1 (MMP1; inters	15.7
421727	Y13163	Hs.107318	SS,Peptidase_M1,	kyurenine 3-monooxygenase (kyurenine 3	15.3
411658	W20227	Hs.22439	hemopexin,Peptidase_M10,SS,Peptidase_M10	ESTs	13.9
400289	X07820	Hs.2258	SS,DENN	matrix metalloproteinase 10 (MMP10; str	13.5
443348	AW673558	Hs.182278	SS,DENN	coiled-coil 2 (phosphorylase kinase, delt	13.0
424086	AI361010	Hs.102267	SS,Lysyl_oxidase	lysyl oxidase	12.8
400256	W72838	Hs.2533	SS	aldehyde dehydrogenase 9 family, member	12.7
406771	AW732575	Hs.47584	TM,K_tetra,ion_trans,SS,TM,K_tetra,ion_1	potassium voltage-gated channel, delayed	11.9
421155	H67879	Hs.102267	SS,Lysyl_oxidase,Aldose_oxim,Epicimerase,S	lysyl oxidase	11.8
424905	NM_002497	Hs.153704	kininase,SS,TM,kininase,polyperenyl_Lynt,	NIMA (never-in mitosis gene a)-related k	11.7
438167	R28363	Hs.24286	SS,TM,7bn_1,p450,rm	ESTs	11.5
456953	AI007673		kininase,	gbtL-BT152-080399-004 BT152 Homo sapien	11.5
423945	AA410943		death,ZU5,TM,Actin_recp,kinase,	gbz32h03.r1 Soares ovary tumor NBtOH H	11.4
445263	H57846	Hs.42586	SS,Acyltransferase,	KIAA1560 protein	11.2

	423887	AL080207	Hs.134685	SS,TM,BRCT,ank,ABC_tran,ABC_tran	DKFZP434G232 protein	10.9
	429569	NM_007050	Hs.225962	SS,TMAY_phosphatase,MAM,fn3,	protein tyrosine phosphatase, receptor t	10.4
	428368	BE440042	Hs.83326	SS,Peptidase_M10,hemopexin,SS,Peptidase	matrix metalloproteinase 3 (stromelysin	10.3
	418912	NM_003685	Hs.99472	SS,TM,Tm_1,SS,TM,7tm_1,	angiotensin receptor 1	10.3
5	451952	AL120173	Hs.301663	SS,pkinase,	ESTs	10.3
	402408	NA		SS,carb_anhydrase	NM_030920-Homo sapiens hypothalamic pro	9.8
	415539	A1733681	Hs.72472	death,ZU5,TM,Activin_rec,pkinase,	BMP-R1B	9.4
	406687	M51126	Hs.272620	SS,Peptidase_M10,hemopexin,SS,Peptidase	matrix metalloproteinase 11 (MMP11; stro	9.1
	419948	AB041035	Hs.93347	Ferric_reduct,TM,Ferric_reduct,	NM_016931-Homo sapiens NADPH oxidase 4 (	9.1
10	400285	NA		TM,ABC_tran,ABC_membrane,	Eos Control	8.8
	400380	AF123050	Hs.44532	SS,TM,ubiquitin,7tm_3,ANF_receptor,ausli	dubiquitin	8.6
	409203	AA780473	Hs.887	SS,p450,SS,p450	cytochrome P450, subfamily IVB, polypept	8.3
	424502	NM_003968	Hs.153887	SS,SS	inositol polyphosphate 4-phosphatase, ty	8.3
	451725	X55724	Hs.2639	SS,Cys_knot,SS	Norrie disease (pseudoglioma)	8.3
15	418092	R45154	Hs.105804	death,ZU5,pkinase,Activin_rec,	ESTs	8.3
	427811	M81057	Hs.130884	SS,Zn_ceroid,Prop,Prop_M14,SS,Prop_M14	glutamine-fructose-6-phosphate transamin	8.2
	423554	M90516	Hs.1574	GATase_2,SS,	ESTs	8.1
	426261	AW242243	Hs.186670	SS,TM,WD40,ubiquitin,E1-E2,ATPass,Calton	peroxisomal farnesylated protein	7.8
20	413374	NM_001034	Hs.75319	SS	ribonucleotide reductase M2 polypeptide	7.6
	432677	NM_004482	Hs.278611	SS,TM,Glycos_transf_2,Ricin_B_lectin,DPP	UDP-N-acetyl-alpha-D-galactosamine:poly	7.6
	456896	D38299	Hs.170917	SS,TM,7tm_1,	prostaglandin E receptor 3 (subtype EP3)	7.5
	453331	A1240655	Hs.8395	SS,TM,disintegrin,Pep_M12B_propept,Reprol	ESTs	7.3
	407721	Y12735	Hs.38018	pkinase,	dual-specificity tyrosine-(Y)-phosphoryl	7.2
	418004	U37519	Hs.87539	SS,aldehyd,SS,aldehyd,	aldehyde dehydrogenase 3 family, member	7.2
25	410555	U92649	Hs.64311	TM,disintegrin,Reprolysin,	a disintegrin and metalloproteinase doma	7.1
	443685	AW204099	Hs.337720		ESTs, Weakly similar to AF128760 1 retin	6.9
	423545	AF000692	Hs.129781	GAF,PDEase	chromosome 21 open reading frame 5	6.8
	401045	C11001883	g[15753278]re	ICE_p20,SS,ICE_p10,ICE_p20,ICE_p10,ICE_p		6.7
30	442082	R41823	Hs.7413	TM,EBP1_bcl,pkinase,SAM,fn3,	ESTs; calyntenin-2	6.7
	413836	A1655499	Hs.161712	TM,Activin_rec,pkinase,death,ZU5,	ESTs	6.6
	404091	NA		TM,7tm_3,ANF_receptor,	Target Exon	6.6
	450865	A1248013	Hs.105532	zf-C2H2	ESTs, Weakly similar to U38588 reverse t	6.6
35	424085	NM_002914	Hs.139226	SS,AAA,Viral_helicase1,rm,	replication factor C (activator 1) 2 (40	6.5
	449650	AF055575	Hs.22838	TM,ion_trans,SS,TM,ion_trans,	calcium channel, voltage-dependent, L ty	6.5
	432304	AA932186	Hs.68297	TM,7tm_1,	ESTs	6.2
	415689	NM_005025	Hs.73569	SS,serpin,	serine (or cysteine) proteinase inhibito	6.2
	442842	AW167087	Hs.131582	SS,lg,Sema,pkinase,	ESTs	6.2
40	428795	R45503	Hs.97489	SS,TM	ESTs, Highly similar to A39769 N-acetyl	6.1
	428479	Y03272	Hs.184572	SS,pkinase,pkinase	cell division cycle 2, G1 to S and G2 to	6.1
	400300	X03353		SS,TM,pkinase,Recep_L_domain,SH2,PH,Furi	HER2 receptor tyrosine kinase (c-erb-b2,	6.1
	447359	NM_012093	Hs.16286	SS,adonylektinase,	adenylate kinase 5	6.1
	402230	NA		SS,TM,p450,	Target Exon	6.1
45	424687	J05070	Hs.151738	SS,Peptidase_M10,fn2,hemopexin,SS,TM,Pep	matrix metalloproteinase 9 (gelatinase B	5.8
	433238	A1572739	Hs.195471	6PFZL,PGAM,	6-phosphofructo-2-kinase/fructose-2,6-bi	5.8
	432239	X81334	Hs.2836	SS,Peptidase_M10,hemopexin,SS,Peptidase	matrix metalloproteinase 13 (collagenase)	5.6
	400286	NA		SS,TM,ABC_tran,ABC_membrane,SS	C16000922.g[1498103]p[17120903] hypothe	5.6
	425247	NM_005940	Hs.155324	SS,Peptidase_M10,hemopexin,SS	matrix metalloproteinase 11 (MMP11; stro	5.4
50	434757	AA828246	Hs.291884	UCH-L1,pkinase,OPR,Rhodanase,AMP-binding,	ESTs	5.4
	438310	AF086120	Hs.102793	SS,TM,UDPGT,caslin_kappa	G protein-coupled receptor 34	5.2
	441111	A1806867	Hs.126594	SS,TM,Phosphodiast,	ESTs	5.2
	452355	M54525	Hs.28202	SS,TM,7tm_1,TM	solute carrier family 25 (mitochondrial	5.2
	427711	K11659	Hs.180408	SS	ATP-binding cassette transporter MRP8	5.1
55	429353	AL117408	Hs.200102	SS,TM,ABC_tran,ABC_membrane,		
	425325	X52730	Hs.1892	SS,NNMT_PNMT_TEMT,SS,NNMT_PNMT_TEMT	STAR	
	448706	AW291095	Hs.21814	phenylethanolamine N-methyltransferase	5.1	
	400563	NA		SS,TM,pkinase,	interleukin 20 receptor, alpha	5.1
60	432777	AA564991	Hs.269477	CIDE-N_pkinase	Target Exon	5.1
	446232	A1281848	Hs.194691	alpha-amylose,	ESTs	5.0
	411165	NM_000169	Hs.68089	SS,TM,7tm_3,Ribosomal_L13	gelic acid induced 3	4.9
	414117	W88559	Hs.1787	SS,Multikase,BTK,pkinase,SH2,SH3,Ribo	galactosidase, alpha	4.9
	411096	U80034	Hs.68583	TM,ion_trans,K_kelra,	proteolipid protein 1 (Polizzeau-Marzbac	4.9
	450506	NM_004460	Hs.418	Peptidase_M3,	mitochondrial isubmodia peptide	4.9
65	417975	AA641836	Hs.30085	SS,DPPIV_N_term,Peptidase_SS,SS,DPPIV	N_fibroblast activation protein, alpha	4.9
	447752	M73700	Hs.105938	SS,tyrosin	hypothetical protein FLJ23186	4.9
	427122	AW057738	Hs.323910	SS,transferrin,7tm_1,transferrin,	lactotransferrin	4.8
				SS,TM,pkinase,Recep_L_domain,SH2,PH,Furi	HER2 receptor tyrosine kinase (c-erb-b2,	4.8



400181	NA	SS,TM,3Beta_HSD,	ENSP00000171555.CDNA FLJ10727 fs, clone	4.6
452093	AA447453	SS,TM,7tm_1,	Homo sapiens mRNA; cDNA DKFZp586M0723 (f	4.6
435542	AA687378	SS,pkinase,RhoGEF,Ig,PH,SH3,	ESTs	4.6
417576	AA339449	AIRS,formyl_transf,GARS,SS,GARS,AIRS,lor	phosphoribosylhydramide formyltransfer	4.6
446089	A860021	pinase	ESTs, Moderately similar to A47582 B-cell	4.6
424420	B5514743	SS,TM,MAPEG,	prostaglandin E synthase	4.5
452190	H28735	SS,TM,PH,SH2,Furin-like,pkinase,Recep_L	Homo sapiens clone PP1498 unknown mRNA	4.5
419886	A345465	pkinase,OPP,	GA-binding protein transcription factor,	4.5
421882	A310275	SS,trefol,SS,TM,Idl_recep_L,SRCR,tyrps	trefol factor 1 (p52)	4.5
446733	AA683380	SS,TM,p450,	ESTs, Weakly similar to fatty acid omega	4.4
453060	AW294092	SS,ras_Y_phosphatase,ras	hypothetical protein MGC16754	4.4
400205	NA		NM_006265*Homo sapiens RAD21 (S. pombe)	4.4
420854	AW296527	SS,TM,Peptidase_M1,	gb U1-H-BWD-ajc-c-07-0-U1.1 NCLGAP_Su	4.3
422890	AF181490	SS,pyr_redox,SS,Ribosomal_L39	prenylcysteine lyase	4.3
441560	F13366	pinase,	Homo sapiens clone Z3736 mRNA sequence	4.3
416445	AL043004	SS,pkinase,	KIAA0135 protein	4.3
439024	R96866	SS,TM,tyrps,vwd,Ig	ESTs	4.3
432882	NM_013257	pinase,pkinase_C,	serum/glucocorticoid regulated kinase-II	4.2
447754	AW073310	pinase,	Homo sapiens cDNA FLJ14142 fs, clone MA	4.2
453775	NM_002910	SS,AAA_P13_P4_kinase,P13Ka,P13K_fdb,P13	replication factor C (eukaryot 1) 4 (37	4.1
431667	A345227	SS,TM,pkinase	ESTs, Weakly similar to B34087 hypotheti	4.1
427899	AA82268	SS,AAA_proteins,ABC_mambrane,ABC_tran,	serum amyloid A1	4.1
422938	NM_001809	SS,TM,hiclose,	centromere protein A (17KD)	4.1
418478	U38945	ank,ank	cyclin-dependent kinase inhibitor 2A (me	4.1
406922	870284	SS,TM,Desaturase,SS	gb stearyl-CoA desaturase [human, adip	4.1
438285	AL133816	SS,Ig,pkinase,LRRT,LRRC,	hypothetical protein FLJ20093	4.1
429922	Z97630	SS,TM,linck_histone,7tm_1	H1 histone family, member 0	4.1
420139	NM_006357	SS,TM,p450,	lipase, hormone-sensitive	4.0
425071	NM_013989	SS,T4_deidnase,T4_deidnase,	deidnase, iodotyrosine, type II	4.0
424511	BE300512	SS,Y_phosphatase,Band_41	ESTs, Moderately similar to ALU7_HUMAN A	4.0
428722	U75456	SS,TIMP,	tissue inhibitor of metalloproteinase 3	3.9
428330	L22524	SS,Peptidase_M10,SS,Peptidase_M10,hemopexin	matrix metalloproteinase 7 (MMP7); uterin	3.9
414631	M31156	SS,cNMP_binding,Rita,HMG_box	protein kinase, cAMP-dependent, regulato	3.9
413589	AW452831	SS,abhydrolase,	ESTs, Highly similar to A4157833 1 nmcl	3.8
418622	BE247559	SS,SH2,P13,SS,TM,PH,SH2,Furin-like,pkinas	growth factor receptor-bound protein 7	3.8
413453	AA128640	SS,Peptidase_C1,gph	ESTs	3.8
403943		p450,SS,p450	C600355:gl 4503225(re NP_000785.1) cyt	3.8
444618	AV853785	SS,TM,7tm_1,	ELL-RELATED RNA POLYMERASE II, ELONGATIO	3.8
40781	AA057264	SS,TM,7tm_1,	ESTs, Weakly similar to (define not av	3.8
427809	M26380	SS,pkase,PLAT,Sec7,PH,	flpocprotein lipase	3.7
418203	X54942	CKS,SS,CKS,	GDC28 protein kinase 2	3.7
454071	AL041793	TM,7tm_1,	ESTs	3.7
424676	Y08565	Glycos_transf_2,Ricin_B_lectin,SS,Glycos	UDP-N-acetyl-alpha-D-galactosamine:poly	3.7
457485	AW301344	SS,Phycoerythrin,Sulfatase	DNA replication factor	3.7
417601	NM_014735	PHD,pkinase,SS	KAA0215 gene product	3.7
446819	AU078943	SS,TM,ehand,Ion_trans	secreted phosphoprotein 1 (osteopontin,	3.7
406625	Y13647	SS,TM,Desaturase,SS	stearyl-CoA desaturase (delta-9-desatur	3.6
428709	AW207475	SS,TM,1_SPRY,	ESTs	3.6
425310	NM_003091	SS,pkinase,vwa,vwa,Glyco_transf_8	neuropeptide Y receptor Y1	3.6
417531	NM_003157	SS,PH,Jadcamase_B	serine/threonine kinase 2	3.6
444781	NM_014400	SS,SH,Jadcamase_B	GPI-anchored metastasis-associated prote	3.6
431493	A781493	SS,PH,p450,p450	ESTs, Weakly similar to A36036 cytochrom	3.6
428966	AF059214	TM,hemopexin,Peptidase_M10,hemopexin,Pep	cholesterol 25-hydroxylase	3.6
414175	A1308876	TM,Ion_trans,K_Jeila,	gb CV4-NN0039-290300-154-106 NN0039 Homo	3.6
455325	AW85978		a disintegrin and metalloproteinase domain	3.6
429597	NM_003816	AoyL_transf,adh_zinc,ketozoyl-synt,pp-bl	fatty acid synthase	3.5
425320	U23544	IMS,SS	polymerase (DNA directed) Iota	3.5
431854	AA383560	SS,TM,SH2,PH,pkinase,Recep_L_domain,Furin-erb-b2	avian erythroblastic leukemia v	3.5
407104	867296	SS	HER2 receptor tyrosine kinase (c-erb-b2,	3.5
449051	AW961400	pkinase,PCLO_box,SS,pkinase,PCLO_box,	serum-inducible kinase	3.5
434398	AA121098	SS,Ig,pkinase,LRRT,LRRC,	hypothetical protein FLJ20093	3.5
454042	DE2570	SS,TM,7tm_1,	ubiquitin specific protease 18	3.5
407786	AA118021	Peptidase_M24,	histone deacetylase 9	3.5
401464	A7303241	DSPc,Mycosin_tail,	dual specificity phosphatase 10	3.5
412970	AB026436	SS,adenylylase,kinase,	adenylylase kinase 5	3.5
412049	N53437	SS,TM,7tm_2,SS,TM,7tm_2	parathyroid hormone receptor 2	3.4
425776	U25128			3.5

	407846	AA426202	Hs.40403	TM,ABC_membrane,ABC_tran,Ribosomal,S4eCbp300-interacting transactivator, wt	3.4
	426293	L34041	Hs.8739	SS,TM,transport_prot,SWIB,RhoGAP,DAG_PE_glycerol-3-phosphate dehydrogenase 1 (so	3.4
	445873	AA256970	Hs.251946	SS,rm,PABP,pkinase,14-3-3,rm	3.4
	418054	NM_002318	Hs.83354	SS,TM,milo_carr,Lysyl oxidase	3.4
5	406815	AA833930	Hs.288036	SS,JIPPT,	3.4
	410530	M25809	Hs.64173	ATP_synth_ab,SS,7tm_1,ATP_synth_ab	3.4
	407021	U52077		SS,p450,SS	3.4
	421198	AF182277	Hs.330780	SS	3.4
	431473	AA825686	Hs.321176	pkinase,	3.4
10	408101	AW865504	Hs.123073	SS,peroxidase,PLAT,	3.4
	422093	NM_001141	Hs.111256	SS,sushi,tyrosin,wwa,rm,flthrdogen_C,fn	3.3
	411393	AW797437	Hs.69771	SS,Peptidase_S8,P,Peptidase_S8,P	3.3
	435767	H73505	Hs.117374	SS,Peptidase_S8,P,Peptidase_S8,P	3.3
	433068	NM_008488	Hs.286215	SS,Peptidase_S8,P,Peptidase_S8,P	3.3
15	426928	AF037082	Hs.172914	SS,adh_short,TGF-beta,7GFb_propeptide	3.3
	414375	H11257	Hs.22568	SS,pkinase,fg,	3.3
	445941	AB267371	Hs.172336	SS,SS,peroxidase,PLAT	3.3
	444542	AI161293	Hs.280380	SS,SS,Peptidase_M1,EGF_fg,lecln_c,sushi	3.3
20	425741	AF042152	Hs.159412	pkinase,	3.3
	434226	Z42047	Hs.263978	SS,TM,7tm_1	3.3
	433264	D85782	Hs.3229	SS,Peptidase_M1,	3.3
	400419	AF084545		TM,Integrin_B,Ricin_B,lecln,rm	3.3
	439750	AL369053	Hs.57884	death,ZU5,pkinase,Activin_rec,	3.3
25	417757	R19897	Hs.106804	SS,TM,7tm_3,ANF_receptor,sushi	3.3
	452194	AI694413	Hs.332649	SS	3.2
	421458	NM_003954	Hs.104576	SS,PC1,RasGEF,hormone_rec,zf-C4,	3.2
	443767	BE562136	Hs.9736	peroxidase,LRRGT,	3.2
	422946	D86983	Hs.115893	SS,p450,p450	3.2
30	423451	AA325602		SS,TM,hehase,	3.2
	451254	AF082235		SS,TM,Activin_rec,pkinase	3.2
	452110	T47657	Hs.28005	TM,p450,eb,	3.2
	439563	AW247529	Hs.6793	SS,DEAD,HRDC,helicase_C,	3.2
	453941	U36817	Hs.36820	SS,TM,transport_prot,SWIB,RhoGAP,DAG_PE-	3.1
35	405654	L34041	Hs.8739	3.1	glycerol-3-
				phosphate dehydrogenase 1 (so	
	453487	R31710	Hs.23540	TM,7tm_1,	3.1
	420911	U77413	Hs.100283	ESTs	3.1
	443171	BE281128	Hs.9030	O-linked N-acetylglucosamine (GlcNAc) tr	3.1
	452256	AK000933	Hs.28861	TM,CDI,7tm_1,	3.1
40	432201	AI538613	Hs.288241	SS,TM,tyrosin,SS,TM,trefoll,tyrosin,tref	3.1
	419150	T29618	Hs.89640	TM,pkinase,fn3,	3.1
	444443	AI149286	Hs.85099	SS	3.1
	426283	NM_003937	Hs.169139	SS,alkylolase,	3.1
	450291	BE588452	Hs.5101	SS,pro_tomerase,	3.1
45	450223	AA416204	Hs.241493	SS,TM,7tm_2,GPS	3.1
	424269	AW137691	Hs.199754	ESTs	3.1
	448105	AW591433	Hs.298241	SS,TM,trefoll,tyrosin,trefoll	3.0
	462560	BE077084	Hs.336432	SS,rm,zf-RanBP,pkinase,C2,pkinase_C,DAG ESTs	3.0

TABLE 21A

Table 21 A shows the accession numbers for those pkeys lacking unigenelD's for Table 21.

- 5 For each probeset, we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

10

Pkey: Unique Eco probeset identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

15

Pkey CAT number Accessions

20

420854	197072_1	AW296927 AJ584514 AI263168 AA281079
423431	228162_1	AA326062 AA325758 AW982182
423945	233956_1	AA410643 AW948953 AA334202 AA332882
451264	863988_1	AI766235 R31400 H29082 H23107
455325	1279475_1	AW6895719 N31451 N41451
456207	155078_1	AA193450

TABLE 21B

**Table 21B** shows the genomic positioning for those pkeys lacking unigene ID's and accession numbers in Table 21. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

10	Pkey:	Unique number corresponding to an Eos probeset		
	Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., <i>Nature</i> (1999) 402:489-495.		
	Strand:	Indicates DNA strand from which exons were predicted.		
	NT_position:	Indicates nucleotide positions of predicted exons.		
15				
	Pkey	Ref	Strand	NT_position
	401045	8117819	Plus	90344-90184,91111-91345
	402230	8966312	Minus	29762-29932
20	402408	9786239	Minus	110326-110491
	402578	9884928	Plus	68350-68496
	403393	6882850	Minus	62554-62712,69449-69602
	403943	7711884	Plus	100742-100904,101322-101503
	404091	7634554	Minus	82121-83229

**TABLE 22: 739 GENES UP-REGULATED IN BREAST CANCER  
COMPARED TO NORMAL ADULT BREAST**

Table 22 shows 739 genes up-regulated in breast cancer compared to normal adult breast. These were selected as for Table 19, except that the ratio was greater than or equal to 3.0, the denominator was the 85<sup>th</sup> percentile value for 12 non-malignant breast specimens, and the 96<sup>th</sup> percentile value amongst the 73 breast cancers was greater than or equal 100 units.

Pkey: Unique Eos probe/seq identifier number  
ExAcon: Exemplar Accession number, Genbank accession number  
UnigeneID: Unigene number  
UnigeneTitle: Unigene gene title  
R1: Ratio of 80<sup>th</sup> percentile tumor to 85<sup>th</sup> percentile normal breast tissue

Pkey	ExAcon	UnigeneID	UnigeneTitle	R1
400292	AA250737	Hs.72472	BMP-R1B	51.5
424735	U13875	Hs.272499	short-chain alcohol dehydrogenase family	38.3
400297	A1127076	Hs.334473	hypothetical protein DKFZp564O1278	29.9
431448	AL137517	Hs.334473	hypothetical protein DKFZp564O1278	26.9
451110	AJ555040	Hs.265398	ESTs, Weakly similar to transformation-r	25.8
431211	M86949	Hs.323733	gap junction protein, beta 2, 25kD (conn	23.2
418233	X54542	Hs.63753	CDCC2 protein kinase 2	22.6
407980	A0463309		gbc21201.s1 Soares_fetal_heart_NbH19W	19.8
414648	A3353776	Hs.901	CD48 antigen (B-cell membrane protein)	18.9
446821	AB012113	Hs.18530	small inducible cytokine subfamily A (Cy	18.0
409041	AB033025	Hs.50081	KIAA1159 protein	17.6
412140	AA219691	Hs.73625	RAB6 Interacting, kinesin-like (rabkines	17.6
407824	AA147884	Hs.9812	Homo sapiens cDNA FLJ14388 fig. clone HE	17.1
453160	AL283307	Hs.239884	H2B histone family, member L	17.0
407137	T97307		gbyes53H5.s1 Soares fetal liver spleen	16.1
425892	D90041	Hs.155856	N-acetyltransferase 1 (arylamine N-acety	16.1
438533	AI440286	Hs.170673	ESTs, Weakly similar to T24832 hypotheti	16.0
428227	AA321649	Hs.2248	small inducible cytokine subfamily B (CX	15.5
444342	NM_014398	Hs.10887	similar to lysosome-associated membrane	15.1
422505	AL120862	Hs.124165	programmed cell death 9 (PDCD9)	14.9
430515	AA746503	Hs.283313	ESTs	14.7
417338	H60720	Hs.81892	KIAA0101 gene product	14.4
452744	AI267652	Hs.30504	Homo sapiens mRNA; cDNA DKFZp434E082 (p14.4	14.1
412448	AI769015	Hs.92127	ESTs	14.2
415538	AI733881	Hs.72472	BMP-R1B	14.1
435498	AW640171	Hs.265398	ESTs, Weakly similar to transformation-r	13.8
438209	AL120659	Hs.6111	aryl-hydrocarbon receptor nuclear transi	13.8
402205	NA		NM_008265*Homo sapiens RAD21 (S. pombe)	13.5
430963	AA489732	Hs.154918	ESTs	13.4
415283	AA948033	Hs.130653	ESTs	13.3
451952	AL120173	Hs.301693	ESTs	13.2
449722	BE280074	Hs.23960	cyclin B1	13.2
406685	M18728		gbcHuman nonspecific crossreacting anti	13.0
406690	M29540	Hs.220529	carcinoembryonic antigen-related cell ad	12.8
429925	NM_000789	Hs.226213	cytochrome P450, 51 (lanosterol 14 alpha	12.8
416498	U33632	Hs.75351	potassium channel, subfamily K, member 1	12.7
432379	AA83046	Hs.146133	ESTs	12.5
441377	BE218239	Hs.202656	ESTs	12.5
458207	AA193450		gbczr40e07.s1 Soares_NbHMPu_S1 Homo sapi	12.4
422805	AA439899	Hs.121017	H2A histone family, member A	12.2
407811	AW190902	Hs.40098	cysteine knot superfamily 1, BMP antagon	12.2
407178	AA195651	Hs.104106	ESTs	12.2

	420931	AF044197	Hs.100431	small inducible cytokine B subfamily (Cy	12.1
	421727	Y13153	Hs.107318	kynurenine 3-monooxygenase (kynurenine 3	12.0
	434408	AI031771	Hs.132566	ESTs	12.0
5	446591	H44186	Hs.15456	PDZ domain containing 1	11.9
	431385	BE175633	Hs.11030	membrane-spanning 4-domain, subfamily A	11.8
	443348	AW873506	Hs.182278	calmodulin 2 (phosphorylase kinase, dolt	11.7
	416602	NM_005159	Hs.75389	rel (chicken)-like 2	11.7
	433365	AF026944	Hs.293797	ESTs	11.6
	437836	AA156781	Hs.74170	metallothionein 1E (functional)	11.5
10	412472	AW975398	Hs.298336	ESTs	11.4
	416030	H15281	Hs.21948	ESTs	11.3
	439979	AW600291	Hs.58823	hypothetical protein FLJ10430	11.3
	420757	X78592	Hs.95915	androgen receptor (dihydrotestosterone r	11.3
	411598	BE336854	Hs.70937	H3 histone family, member A	11.2
15	423600	AI633559	Hs.310369	ESTs	11.2
	430770	AA765694	Hs.123295	ESTs	11.0
	421037	AI684608	Hs.187653	programmed cell death 9 (PDCD9)	10.9
	452451	N76223	Hs.108108	transcription factor	10.7
	408269	AA570953	Hs.22972	hypothetical protein FLJ13352	10.6
20	417791	AW965339	Hs.111471	ESTs	10.6
	447258	AI370413	Hs.35653	hypothetical protein FLJ22418	10.4
	424001	W87883	Hs.137476	paternally expressed 10	10.4
	447342	AI199268	Hs.19322	Homo sapiens, similar to RIKEN cDNA 2010	10.4
	424905	NM_002497	Hs.153704	NIMA (never in mitosis gene a)-related k	10.1
25	453619	H87848	Hs.33922	Homo sapiens, clone MGC:9084, mRNA, comp10.1	10.1
	442942	AW167087	Hs.131562	ESTs	10.1
	434377	AW137148	Hs.306593	Homo sapiens cDNA FLJ11362 fls, clone HE	10.1
	427217	AA399272	Hs.144341	ESTs	10.1
	445730	AI624342	Hs.170042	ESTs	10.0
30	432887	AI926047	Hs.162859	ESTs	10.0
	452243	AL355715	Hs.28555	programmed cell death 9	9.9
	424580	AW905398	Hs.46021	hypothetical protein FLJ20086	9.9
	432169	Y00911	Hs.2910	phosphoribosyl pyrophosphate synthetase	9.9
	430650	H23769	Hs.144530	EST	9.9
35	418836	AI656469	Hs.161712	ESTs	9.8
	430291	AV680345	Hs.238126	CGI-49 protein	9.8
	444655	BE613128	Hs.47783	B aggressive lymphoma gene	9.7
	407377	C16391		gbcC16391 Clontech human acrota polyA mRNA	9.7
	445413	AA151342	Hs.12677	CGI-147 protein	9.7
40	443462	AI064890	Hs.171176	ESTs	9.7
	442145	AI022650	Hs.8117	erbB2-interacting protein ERBIN	9.7
	435570	AF212222	Hs.177812	uncharacterized bone marrow protein BM04	9.7
	439820	AL360204	Hs.283853	Homo sapiens mRNA full length insert cDN	9.6
	428936	AF059214	Hs.194887	cholesterol 25-hydroxylase	9.6
45	449448	D60730	Hs.57471	ESTs	9.6
	433929	AI375499	Hs.27379	ESTs	9.5
	432731	R31178	Hs.257820	fibronectin 1	9.3
	411815	AA158879	Hs.125790	leucine-rich repeat-containing 2	9.3
	415385	R17798	Hs.7635	COBIV-like protein	9.3
50	422026	U80736	Hs.110826	tridecadecyl repeat containing 9	9.2
	432596	AI224741	Hs.278481	matrilin 3	9.2
	439451	AF086270	Hs.278554	heterochromatin-like protein 1	9.2
	423945	AA410943		gbcz32h203.1 Soares ovary tumor N6HOT H	9.1
	442432	BE093589	Hs.38178	hypothetical protein FLJ23458	9.1
55	446715	AI337735	Hs.173819	ESTs, Moderately similar to ZNF1_HUMAN Z	9.0
	408771	AW732573	Hs.47564	potassium voltage-gated channel, delayed	9.0
	437021	AI076089	Hs.292239	ESTs	9.0
	428479	Y00272	Hs.184572	cell division cycle 2, G1 to S and G2 to	8.9
	428839	AI767755	Hs.82302	Homo sapiens cDNA FLJ14614 fls, clone NT	8.9
60	402406	NA		NM_005020? Homo sapiens hypothetical pro	8.8
	418601	AA279490	Hs.86368	calmagin	8.8
	428327	W03242	Hs.44868	Homo sapiens clone TCCTA00151 mRNA sequ	8.8
	418519	AI198719	Hs.178376	ESTs	8.8
	440821	AW298024	Hs.150434	ESTs	8.8
65	446142	AI754893	Hs.145668	ESTs	8.8
	418196	AI745849	Hs.26549	KIAA1708 protein	8.7
	447178	AW84641	Hs.192417	ESTs	8.7

	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaph	8.6
	415857	AA868115	Hs.127797	Homo sapiens cDNA FLJ11381 fls, clone H	8.6
	435081	AB51474	Hs.163944	ESTs	8.6
5	431374	BE258532	Hs.251871	CTP synthase	8.4
	417866	AW007603	Hs.82772	collagen, type XI, alpha 1	8.4
	437211	AA362207	Hs.5509	ecotropic viral integration site 2B	8.3
	437751	AA767373	Hs.36689	ESTs, Moderately similar to ALU1_HUMAN A	8.3
	423887	AL080207	Hs.134585	DKFZP434G232 protein	8.2
	440941	BE268362	Hs.7535	COB-W-like protein	8.2
10	429859	NM_007050	Hs.225852	protein tyrosine phosphatase, receptor t	8.2
	410193	AJ132692	Hs.59757	zinc finger protein 281	8.2
	431725	X65724	Hs.2839	Norrie disease (pseudoglioma)	8.1
	446258	AI283476	Hs.263478	ESTs	8.1
	416747	AW876523	Hs.15929	hypothetical protein FLJ12910	8.1
15	434424	AI811202	Hs.325335	Homo sapiens cDNA: FLJ23523 fls, clone L	8.1
	421650	AA781795	Hs.122587	ESTs	8.0
	429534	AW976987	Hs.183327	ESTs, Weakly similar to 2108260A B cell	8.0
	457465	AW301344	Hs.122938	DNA replication factor	8.0
	427951	AW235165	Hs.143134	ESTs	8.0
20	436481	AA739597	Hs.5199	HSPC150 protein similar to ubiquitin-con	8.0
	418216	AA662240	Hs.263099	AF15q14 protein	8.0
	418250	U29925	Hs.83918	adenosine monophosphate deaminase (iso	7.9
	400285	NA		Eos Control	7.9
	401464	AF030241	Hs.8028	histone deacetylase 5	7.9
25	407242	M18728		gbHuman nonspecific crossreacting anti	7.8
	422232	D43945	Hs.113274	transcription factor EG	7.8
	454024	AA983527	Hs.293907	hypothetical protein FLJ23403	7.8
	444542	AI161293	Hs.280380	aminopeptidase	7.8
30	436396	AI683487	Hs.152213	wingless-type MMTV integration site fami	7.7
	437204	AL110216	Hs.12285	ESTs, Weakly similar to I55214 salivary	7.6
	408805	H69912	Hs.48269	vacuolin related kinase 1	7.6
	437207	T27503	Hs.15929	hypothetical protein FLJ12910	7.6
	442816	AK001741	Hs.8739	hypothetical protein FLJ10879	7.6
	426283	NM_003037	Hs.166139	lysine/arginine (L-lysine/arginine hydrolase)	7.5
35	424687	J65070	Hs.151738	matrix metalloproteinase 9 (gelatinase B	7.5
	446315	NM_016293	Hs.14770	bridging Integrator 2	7.5
	433426	H69125	Hs.133625	ESTs	7.5
	406639	M97711		gbHuman T-cell receptor (V beta 18.1, J	7.5
	420077	AW512260	Hs.87767	ESTs	7.4
40	457332	AA961694	Hs.105187	kinesin protein 9 gene	7.4
	422938	NM_001809	Hs.1584	centromere protein A (17kD)	7.4
	447555	AI391962	Hs.180863	Homo sapiens, clone MGC:12316, mRNA, com7.4	7.4
	444618	AV653785	Hs.173334	ELL-RELATED RNA POLYMERASE II, ELONGATIO	7.3
45	410381	BE391804	Hs.62861	guanylate binding protein 1, interferon-	7.3
	400285	NA		NM_003292/Homo sapiens translocated prom	7.3
	439509	AF086332	Hs.58314	ESTs	7.3
	407771	AL138272	Hs.82713	ESTs	7.3
	407202	N58172	Hs.109370	ESTs	7.3
	433086	AU078803	Hs.282975	carboxylesterase 2 (intestine, liver)	7.2
50	422094	AF129635	Hs.272027	F-box only protein 5	7.1
	430832	AI073813	Hs.100856	ESTs, Weakly similar to JE0360 Anterior	7.1
	430287	AW182459	Hs.125759	ESTs, Weakly similar to LEU5_HUMAN LEUKE7.0	7.0
	423739	AA398155	Hs.87800	ESTs	7.0
	448212	AI475858		gbt087007.x1 NCL_CGAP_CLL1 Homo sapiens7.0	7.0
55	407277	AW170035	Hs.328736	Homo sapiens breast cancer antigen NY-BR	7.0
	454440	BE062906	Hs.28338	KIAA1546 protein	7.0
	444783	AK001468	Hs.82180	anillin (Drosophila Scraps homolog), act	7.0
	421373	AA088229	Hs.187771	ESTs	6.9
	431980	AW241821	Hs.301927	c6.1A	6.9
60	424704	AI263293	Hs.152065	cytochrome P450, subfamily 1U (arachido	6.8
	449517	AW60105	Hs.23643	serine/threonine protein kinase MASK	6.8
	436940	AW449211	Hs.105445	GNF family receptor alpha 1	6.8
	414080	AA135257	Hs.47783	B aggressive lymphoma gene	6.8
	441243	AI787055	Hs.193002	ESTs	6.7
65	406380	AF123050	Hs.44532	diubiquitin	6.7
	422956	BE548072	Hs.122579	hypothetical protein FLJ10481	6.7
	446551	AA393907	Hs.87179	ESTs	6.7

	419839	U24577	Hs.53304	phospholipase A2, group VII (platelet-ac	6.7
	437740	AA810265	Hs.122915	ESTs	6.7
	421582	A1910275	Hs.1406	trefol factor 1 (p52)	6.7
5	427356	AW023482	Hs.57949	ESTs	6.6
	429597	NM_003818	Hs.2442	a disintegrin and metalloproteinase dom	6.6
	422634	NM_018010	Hs.119821	CGI-62 protein	6.6
	421072	A215039	Hs.89113	ESTs	6.5
	427716	A178680	Hs.25933	ESTs	6.5
10	411000	N04449	Hs.201619	ESTs, Weakly similar to S38583 SEB48 p	6.5
	449343	A151418	Hs.272458	protein phosphatase 3 (formerly 2B), ca	6.4
	409757	NM_001889	Hs.123114	cystatin SN	6.4
	447164	AF028941	Hs.17518	Homo sapiens cigs mRNA, partial sequenc	6.4
	450936	X52509	Hs.161640	tyrosine aminotransferase	6.4
	418848	A1820861	Hs.193485	ESTs	6.4
15	424902	NM_003866	Hs.153687	Inositol polyphosphate 4-phosphatase, ty	6.4
	452838	U65011	Hs.30743	preferentially expressed antigen in mela	6.4
	439452	AA918317	Hs.57987	B-cell CLL/lymphoma 11B (zinc finger pro	6.4
	407286	AJ235664		gb/Homo sapiens mRNA for immunoglobulin	6.3
20	411078	AJ222020	Hs.182364	CocosaChip	6.3
	433001	AF217513	Hs.279505	clone HQ6310 PROQ310p1	6.3
	434340	A1153043	Hs.122655	ESTs, Weakly similar to T17226 hypothel	6.2
	425903	AA394183	Hs.26873	ESTs	6.2
	402578		C1001134.g2[117372]p[1]65981 fatty ac	6.2	
	409846	AW161391	Hs.709	deoxycytidine kinase	6.1
25	430447	W17064	Hs.332848	SWI/SNF related, matrix associated, acti	6.1
	432415	T16971	Hs.289014	ESTs, Weakly similar to A43932 mucin 2 p	6.1
	443709	A1082892	Hs.134652	ESTs	6.1
	420929	A1894143	Hs.286251	programmed cell death 4	6.1
30	428368	BE440042	Hs.83326	matrix metalloproteinase 3 (stromelysin	6.1
	428248	A1126772	Hs.40479	ESTs	6.0
	402344	BE463721	Hs.97101	putative G protein-coupled receptor	6.0
	453392	U23732	Hs.32964	SRY (sex determining region Y)-box 11	6.0
	425397	J04086	Hs.155346	topoisomerase (DNA) II alpha (170kd)	6.0
	418037	H13509	Hs.83169	matrix metalloproteinase 1 (MMP1; inters	6.0
35	426595	AB007083	Hs.165140	KIAA0403 protein	6.0
	437608	AA761805	Hs.292308	ESTs, Weakly similar to ALU1_HUMAN ALU	6.0
	427408	AA583206	Hs.2155	RAR-related orphan receptor A	6.0
	406687	M31128	Hs.272820	matrix metalloproteinase 11 (MMP11; stro	6.0
	418092	R45154	Hs.106804	ESTs	6.0
40	447051	AW139130	Hs.160951	ESTs, Weakly similar to Con1 [H.sapiens]	6.0
	441233	AA972865	Hs.135686	ESTs	6.0
	432239	X81334	Hs.2836	matrix metalloproteinase 13 (collagenase	6.0
	435106	AA100847	Hs.193380	ESTs, Highly similar to AF174800 1 F-box	5.9
	435525	A1631297	Hs.123310	ESTs	5.9
45	458809	AW972512	Hs.20985	sn3-associated polypeptide, 30kD	5.9
	410785	AW803341		gbJL2-UM0079-060500-050-D03 UM0079 Homo5.9	5.9
	422576	BE548555	Hs.118554	CGI-83 protein	5.9
	451336	A1793124	Hs.144479	ESTs	5.9
	441881	AW568904	Hs.179566	hypothetical protein FLJ22624	5.8
50	412022	A1035043	Hs.24143	Wiskott-Aldrich syndrome protein Interac	5.8
	416636	N32536	Hs.42645	solute carrier family 16 (monocarboxylic	5.8
	447350	A1375572	Hs.172634	ESTs	5.8
	434094	AA305599	Hs.238205	hypothetical protein PRO2013	5.8
	409151	AA308105	Hs.50785	SEC22, vesicle trafficking protein (S. c	5.8
55	448807	A1571940	Hs.7549	ESTs	5.8
	452281	T93500	Hs.28792	Homo sapiens cDNA FLJ11041 fls, clone PL	5.8
	421281	AJ299139	Hs.17517	ESTs	5.8
	430361	A1033965	Hs.239206	sterol-C4-methyl oxidase-like	5.8
60	400289	X07820	Hs.2258	matrix metalloproteinase 10 (MMP10; str	5.7
	440527	AV657117	Hs.184184	ESTs, Moderately similar to S65857 alpha	5.7
	434674	AA631879	Hs.136085	ESTs	5.7
	426320	W47595	Hs.169300	transforming growth factor, beta 2	5.7
	452401	NM_007115	Hs.78352	tumor necrosis factor, alpha-induced pro	5.7
	446683	BE514699	Hs.108823	hypothetical protein MGC147597	5.7
65	438199	AW016531	Hs.122147	ESTs	5.7
	446203	Z47563	Hs.14286	flavin containing monooxygenase 5	5.7
	428336	AA503115	Hs.183752	microseminoprotein, beta-	5.6



	430379	AF134149	Hs.240395	potassium channel, subfamily K, member 6	5.6
	422535	BE218705	Hs.121378	metallothionein-like 5, testis-specific	5.6
	444758	ALJ44878	Hs.118599	3-hydroxy-3-methylglutaryl-Coenzyme A re	5.6
	434326	AF038159	Hs.9329	chromosome 20 open reading frame 1	5.6
5	400301	X03835	Hs.1657	estrogen receptor 1	5.6
	447078	AW885727	Hs.301570	ESTs	5.6
	432015	AL157504	Hs.158115	Homo sapiens mRNA; cDNA DKFZp586C0724 (f5.5	5.5
	438691	AA095288	Hs.212184	ESTs	5.5
	439809	R41398	Hs.101774	hypothetical protein FLJ23045	5.5
10	415768	AW418196	Hs.257824	hypothetical protein FLJ13782	5.5
	465373	BE247706	Hs.89751	membrane-spanning 4-domains, subfamily A	5.5
	401645	NA	C16001.440*gl 12330704 gb AAG52890.1 AF36.5		5.5
	437967	BE277414	Hs.5947	mel transforming oncogene (derived from	5.5
	448885	AI734009	Hs.127699	KIAA1603 protein	5.4
15	439138	AI742605	Hs.193698	ESTs	5.4
	440270	NM_015986	Hs.7120	cytokine receptor-like molecule 9	5.4
	437536	X81221	Hs.144488	ESTs	5.4
	438167	R26363	Hs.24286	ESTs	5.4
	452741	BE329214	Hs.30503	Homo sapiens cDNA FLJ11344 f8, clone PL	5.4
20	426214	H59846	Hs.128356	ESTs, Moderately similar to ALLU7_HUMAN A	5.4
	413554	AA319146	Hs.75426	secretogranin II (chromogranin C)	5.4
	422867	L32137	Hs.1584	cartilage oligomeric matrix protein (COM	5.4
	434263	N34895	Hs.44648	ESTs	5.4
	446382	AW205188	Hs.160823	ESTs	5.4
25	422406	AF025441	Hs.116206	QIA-interacting protein 5	5.3
	438321	AA576635	Hs.6183	CGI-48 protein	5.3
	418310	AA814100	Hs.86693	ESTs	5.3
	419825	U91816	Hs.91640	nuclear factor of kappa light polypeptid	5.3
	450701	H39560	Hs.289467	Homo sapiens cDNA FLJ12280 f5, clone MA	5.3
30	445800	AF070626	Hs.13429	Homo sapiens clone 24787 mRNA sequence	5.2
	449051	AW951400	Hs.333528	HER2 receptor tyrosine kinase (c-erb-b2,	5.2
	418758	AW959311	Hs.172012	hypothetical protein DKFZp434J037	5.2
	431070	AW408164	Hs.249104	transcription factor 19 (SC1)	5.2
	417079	U65590	Hs.81134	interleukin 1 receptor antagonist	5.2
35	421928	AF013758	Hs.109643	polyadenylate binding protein-interactin	5.2
	428894	AK000713	Hs.193736	hypothetical protein FLJ20705	5.2
	427427	AF077345	Hs.177936	ESTs	5.2
	403485		C3001813*gl 12737279 ref XP_012183.1	k	5.2
	422168	AA588894	Hs.112408	S100 calcium-binding protein A7 (psorias	5.1
40	421937	AI878857	Hs.106708	hematological and neurological expressed	5.1
	426752	X88490	Hs.172004	titin	5.1
	453310	X70697	Hs.553	solute carrier family 6 (neurotransmitter	5.1
	423189	M81933	Hs.1634	cell division cycle 25A	5.1
	412281	AI810054	Hs.14119	ESTs	5.1
45	447513	AW956776	Hs.313500	ESTs, Moderately similar to ALLU7_HUMAN A	5.1
	453031	AL121276	Hs.25144	ESTs	5.1
	404347			Target Exon	5.1
	431806	M30703	Hs.270833	amphiregulin (schwannoma-derived growth	5.1
50	429113	D28235	Hs.196384	prostaglandin-endoperoxide synthase 2 (p	5.1
	436291	BE588452	Hs.5101	protein regulator of cytokinesis 2	5.1
	450803	R43646	Hs.12422	ESTs	5.1
	434725	AK000796	Hs.4104	hypothetical protein	5.0
	436981	H74319	Hs.188620	ESTs	5.0
	407376	AA963138	Hs.142287	ESTs, Weakly similar to ALLU7_HUMAN III	5.0
55	431689	AA305688	Hs.267695	UDP-Gal4betaGlcNAc beta 1,3-galactosyltr	5.0
	405348	NA	C7001634*gl 12698061 db BAB21849.1	AB	5.0
	436196	AK001084	Hs.333498	Homo sapiens cDNA FLJ10222 f5, clone HE	5.0
	437065	ALD36450	Hs.103238	ESTs	5.0
	419196	AI969442	Hs.59638	hypothetical protein FLJ10808	5.0
60	429412	NM_002025	Hs.2407	POU domain, class 2, associating factor	5.0
	446519	ALU076543	Hs.3113	secreted phosphoprotein 1 (osteopontin,	4.9
	403329	NA		Target Exon	4.9
	442875	BE623003	Hs.23626	Homo sapiens clone TCCCTA00142 mRNA sequ	4.9
	442441	AI820662	Hs.129698	ESTs	4.9
65	430375	AW371048	Hs.93766	H4 histone family, member H	4.9
	424126	AW96163		gb EST378236 MAGS resequences, MGI Homol	4.9
	408873	ALD45017	Hs.182278	catmodulin 2 (phosphorylase kinase, delt	4.9

	407910	AA502274	Hs.41296	fibronectin leucine rich transmembrane p	4.9
	432606	NM_002104	Hs.3056	granzyme K (serine protease, granzyme 3)	4.9
	453204	R10789	Hs.191990	ESTs	4.8
5	462020	AA722012	Hs.255757	ESTs, Weakly similar to AT2A_HUMAN POTEN4.8	4.8
	449048	Z45051	Hs.22520	similar to S58401 (cattle) glucose induc	4.8
	408389	R38438	Hs.182575	solute carrier family 15 (H777) transport	4.8
	431645	AF078949	Hs.256483	dynamin light chain-4	4.8
	423575	C18863	Hs.163443	Homo sapiens cDNA FLJ11576 fs, clone HE	4.8
	444248	H83281	Hs.10710	hypothetical protein FLJ20417	4.8
10	421524	AA312082	Hs.105445	GNF family receptor alpha 1	4.8
	452827	A1571835	Hs.55468	ESTs	4.8
	414222	AL135173	Hs.878	sorbitol dehydrogenase	4.8
	460806	AL161989	Hs.77324	eukaryotic translation termination facto	4.8
15	419078	M93119	Hs.89584	insulinoma-associated 1	4.8
	418973	AA233056	Hs.191518	ESTs	4.8
	447033	A1357412	Hs.157601	ESTs	4.8
	451621	A1879148	Hs.26770	fatty acid binding protein 7, brain	4.7
	419968	X04430	Hs.53913	interleukin 6 (interferon, beta 2)	4.7
20	424326	NM_014479	Hs.145296	dihistidinyl protease	4.7
	431585	BE242803	Hs.252823	hypothetical protein FLJ10326	4.7
	429284	AA085974	Hs.198793	Homo sapiens cDNA: FLJ124633 fs, clone H	4.7
	416814	AW152307	Hs.80042	dolichyl-P-Glc:Man6GlcNAc2-PP-dolichylgl	4.7
	439887	NM_015310	Hs.6763	KIAA0942 protein	4.7
	429667	A1875749	Hs.211508	nucleoporin 153Kd	4.7
25	422880	AF228704	Hs.121524	glutathione reductase	4.6
	405801		NM_000360	Homo sapiens choriolidermia (Ra4.6	4.6
	432435	BE218886	Hs.282070	ESTs	4.6
	439544	W26354	Hs.28891	hypothetical protein FLJ11360; antisense p	4.6
30	425354	U82027	Hs.155935	complement component 3a receptor 1	4.6
	436027	A1864053	Hs.39972	ESTs, Weakly similar to I38588 reverse t	4.6
	424823	AW963062	Hs.337404	ESTs	4.6
	403366	NA		Target Exon	4.6
	402542			Target Exon	4.6
35	450193	A1818071	Hs.15807	Homo sapiens Fanconi anemia complementa	4.6
	411678	A1807114	Hs.71465	squalene epoxidase	4.6
	455844	A1264155	Hs.152981	CDP-diacylglycerol synthase (phosphatida	4.6
	448072	A1459306	Hs.24908	ESTs	4.5
	408045	AW138959	Hs.245123	ESTs	4.5
	423782	A1472209	Hs.323117	ESTs	4.5
40	447388	AW630534	Hs.76277	Homo sapiens, clone MGC:8381, mRNA, comp4.5	4.5
	448140	AF146761	Hs.20450	BCM-like membrane protein precursor	4.5
	452561	A1892181	Hs.49169	KIAA1634 protein	4.5
	425331	AW962128		gb:EST374201 MAGE resequences, MAG Hom4.5	4.5
45	428801	AW277121	Hs.254881	ESTs	4.5
	428500	A1815395	Hs.184641	fatty acid desaturase 2	4.5
	426075	AW513091	Hs.270149	ESTs, Weakly similar to 2109260A B cell	4.4
	437259	A1377755	Hs.120595	ESTs	4.4
	400409	AF153341	Hs.285954	Homo sapiens winged helix/orkhead trans	4.4
	412883	AA121573	Hs.58757	zinc finger protein 281	4.4
50	428689	A1815206	Hs.99395	ESTs	4.4
	401865			Target Exon	4.4
	418819	AA228776	Hs.191721	ESTs	4.4
	406348			Target Exon	4.4
	412138	AW895387		gb:QV4-NN0038-300300-157-c10 NN0038 Homo4.4	4.4
55	428550	AW287880	Hs.98861	ESTs	4.4
	411743	AW862244		gb:QV4-CT0361-301299-074-b05 CT0361 Homo4.4	4.4
	429986	BE081342	Hs.283037	HSPC039 protein	4.4
	423291	NM_004129	Hs.126580	guanylate cyclase 1, soluble, beta 2	4.4
	423496	AL110151	Hs.128797	DKFZP586D0824 protein	4.4
60	452190	H26735	Hs.91688	Homo sapiens clone PP1458 unknown mRNA.4.4	4.4
	424871	NM_004529	Hs.153595	low density lipoprotein-related protein	4.3
	429575	AA705003	Hs.89387	ESTs	4.3
	429822	Z97630	Hs.226117	H1 histone family, member 0	4.3
	421378	Y15221	Hs.103582	small inducible cytokine subfamily B (C)	4.3
65	400300	X03363		HER2 receptor tyrosine kinase [c-erb-B2,	4.3
	437258	ALD41243	Hs.174104	ESTs	4.3
	446595	T57448	Hs.15467	hypothetical protein FLJ20725	4.3

	403011	NA	ENSP00000215330: Probable serine/threonine	4.3
	419055	A0565384	Hs.11571	4.3
	419851	NM_001914919	Hs.1165	4.3
	407786	AA597538	Hs.38972	4.3
5	429163	AB014604	Hs.167565	4.3
	442914	AW188551	Hs.59519	4.3
	441029	A081795	Hs.179246	4.3
	452184	A084413	Hs.332849	4.3
	414821	M63835	Hs.77424	4.2
10	410102	AW248508	Hs.279727	4.2
	452110	T47687	Hs.28905	4.2
	442007	AA301116	Hs.142838	4.2
	417318	AW953937	Hs.12891	4.2
15	431816	AW510444	Hs.191705	4.2
	443646	A065198	Hs.164226	4.2
	419169	AW851980	Hs.262346	4.2
	448539	BE091928	Hs.16244	4.2
	423242	AL039402	Hs.125763	4.2
	432146	A092553	Hs.308538	4.2
20	406038	T97490	Hs.50002	4.2
	445825	BE245743	Hs.288529	4.2
	425139	AW630488	Hs.325820	4.2
	447387	BE247676	Hs.18442	4.2
	410166	AK001378	Hs.59346	4.1
25	437295	AW775318	Hs.88417	4.1
	430486	BE062108	Hs.241651	4.1
	441790	AW294909	Hs.132208	4.1
	410129	BE244074	Hs.58631	4.1
	427521	AW973392	Hs.290585	4.1
30	425247	NM_008940	Hs.155324	4.1
	412886	AF041163	Hs.74647	4.1
	411163	BE062828	gb:01336534F1 NIH_MGC_44 Homo sapiens c4.1	4.1
	444301	AK000136	Hs.10760	4.1
	425711	AA363471	Hs.180669	4.1
35	406850	NA	Target Exon	4.1
	440283	AI732892	Hs.190489	4.0
	432441	AW292425	Hs.163484	4.0
	400284	NA	estrogen receptor 1	4.0
	417341	N91453	Hs.102987	4.0
40	425732	U20158	Hs.2488	4.0
	411393	AW797437	Hs.68771	4.0
	425704	U79293	Hs.159264	4.0
	416584	AA013051	Hs.91417	4.0
	419092	J05581	Hs.89803	4.0
45	443147	A034351	Hs.18050	4.0
	406633	AW963372	Hs.46677	4.0
	433404	T32582	Hs.102720	4.0
	421505	BE302796	Hs.105067	4.0
	417920	BE250127	Hs.82906	3.9
50	414632	AW630088	Hs.76550	3.9
	413782	AW411479	Hs.846	3.9
	404580	NA	FK506-binding protein 4 (59kD)	3.9
	452046	AB018345	Hs.27657	3.9
	459587	A031966	gb:zk15604.s1 Soares_pregnant_uterus_Nbh3.9	3.9
55	416658	U03272	Hs.79432	3.9
	426647	AA243484	Hs.294101	3.9
	428353	AL117406	Hs.200102	3.9
	419036	AW134924	Hs.180325	3.9
60	416916	X07871	Hs.89476	3.9
	421977	V094197	Hs.110165	3.9
	442567	A021163	Hs.130261	3.9
	421168	AF182277	Hs.330780	3.9
	431701	AW935480	Hs.14658	3.9
	418526	BE019020	Hs.85838	3.9
65	414928	NM_002543	Hs.77729	3.9
	422730	AA808876	Hs.25933	3.9
	419741	NM_007018	Hs.93002	3.9

	430047	AA263172	Hs.35	protein tyrosine phosphatase, non-recept	3.9
	458614	AA488957	Hs.170361	ESTs, Weakly similar to Z195_HUMAN ZINC	3.8
	429514	AW236861	Hs.193139	ESTs	3.8
	434521	NM_002267	Hs.3886	karyopherin alpha 3 (Importin alpha 4)	3.8
5	409426	U40462	Hs.54452	zinc finger protein, subfamily 1A, 1 (R)	3.8
	439580	BE565847	Hs.74899	hypothetical protein FLJ112820	3.8
	424028	AF055094	Hs.153692	Homo sapiens cDNA FLJ14354 fls, clone Y7	3.8
	400021			APFX control - HUMISGF3A/M67935_MA	3.8
	453403	BE468939	Hs.81779	Homo sapiens cDNA FLJ13591 fls, clone PL	3.8
10	445841	A1267371	Hs.172636	ESTs	3.8
	434378	AA831739	Hs.336440	EST	3.8
	428220	AW207206	Hs.136319	ESTs	3.8
	439176	AA46444	Hs.190394	ESTs, Weakly similar to B28096 Tine-1 pr	3.8
	401045			C11001883* <sub>g</sub> [6753278]ref[NP_033938.1] c	3.8
15	430178	AW449612	Hs.152475	ESTs	3.8
	423397	NM_001638	Hs.1652	chemokine (C-C motif) receptor 7	3.8
	447630	A160149	Hs.44885	lymphoid enhancer-binding factor 1	3.8
	436391	AJ227682	Hs.146274	ESTs	3.8
	413011	AW068115	Hs.821	biglycan	3.8
20	422121	A167949	Hs.179833	ESTs	3.8
	452268	NM_003512	Hs.28777	H2A histone family, member L	3.8
	427811	M81057	Hs.180884	carboxypeptidase B1 (liver)	3.8
	415579	AA165232	Hs.222099	ESTs	3.8
	437330	AA133944	Hs.50115	Homo sapiens mRNA; cDNA DKFZp781J112 (f	3.8
25	427122	AW057735	Hs.323910	HER2 receptor tyrosine kinase (c-erb-b2,	3.7
	400286	NA		C16000922 <sub>g</sub> [p]7498103[p]1720503 hypothe	3.7
	420281	A1623693	Hs.191533	ESTs	3.7
	419926	AW900992	Hs.93796	DKFZP586D2223 protein	3.7
	417541	A1952191	Hs.180040	hypothetical protein FLJ22439	3.7
30	426172	AA371307	Hs.123596	ESTs	3.7
	429836	A1619602	Hs.211577	kinase II (kinesin receptor)	3.7
	467001	J03259	Hs.2962	vitamin D (1,25-dihydroxyvitamin D3) re	3.7
	424109	AW406878		gb:U1HF-BL0-actg-g-06-0-U1r1 NIH_MGC_37	3.7
35	417022	NM_014737	Hs.80905	Ras association (RapGDS/AF-6) domain fam	3.7
	436222	A1206737	Hs.122810	Homo sapiens cDNA FLJ11469 fls, clone HE	3.7
	430448	AK335553	Hs.13303	Homo sapiens cDNA: FLJ21784 fls, clone H	3.7
	432729	AK000292	Hs.278732	hypothetical protein FLJ20285	3.7
	413916	NA9813	Hs.75615	apolipoprotein C-II	3.7
	421682	NM_014141	Hs.105552	cell recognition molecule Casp2	3.7
40	441833	AW958544	Hs.112242	normal mucosa of esophagus specific 1	3.7
	403761	AA057264	Hs.238936	ESTs, Weakly similar to (define not ava	3.7
	406153			Target Exon	3.7
	445963	AW878606	Hs.148006	ESTs	3.7
45	453464	AB84811	Hs.32699	receptor (calcitonin) activity modifying	3.7
	448918	AB911152	Hs.22572	KIA04580 protein	3.7
	413326	AF113676	Hs.287861	serine (or cysteine) proteinase inhibitor	3.6
	448069	U75248	Hs.20191	seven in absentia (Drosophila) homolog 2	3.6
	463313	BE005771	Hs.153746	hypothetical protein FLJ22490	3.6
50	425234	AW152226	Hs.165909	ESTs, Weakly similar to 138022 hypotheti	3.6
	419941	X98654	Hs.93837	phosphatidylinositol transfer protein, m	3.6
	402397	AF188625	Hs.189507	phospholipase A2, group IID	3.6
	430378	Z29572	Hs.2556	tumor necrosis factor receptor superfam	3.6
	448106	AB000470	Hs.171941	ESTs	3.6
	426431	NM_000458	Hs.169883	transcription factor 2, hepatic; LF-33;	3.6
55	431843	AA516420	Hs.183526	ESTs, Weakly similar to 138022 hypotheti	3.6
	426878	BE069341		gb:CV3-BT0381-270100-073-c08 BT0381 Homo3.6	3.6
	434081	AW024973	Hs.283675	NP0009 protein	3.6
	445292	AV853264	Hs.133932	Homo sapiens cDNA FLJ14666 fls, clone NT	3.6
	452101	T09398	Hs.10844	Homo sapiens cDNA FLJ14476 fls, clone MA	3.6
60	427381	NM_014768	Hs.179703	KIAA0129 gene product	3.6
	409047	AW961434	Hs.31539	ESTs	3.6
	416820	NM_004024	Hs.80206	glucose-6-phosphate dehydrogenase	3.6
	410386	W26187	Hs.3327	Homo sapiens cDNA: FLJ22219 fls, clone H	3.6
	440516	S42303	Hs.161	cadherin 2, type 1, N-cadherin (neuronal	3.6
65	434380	AW015415	Hs.127780	ESTs	3.6
	429970	BE278891	Hs.194691	retinoic acid induced 3	3.6
	415079	R43179	Hs.22895	hypothetical protein FLJ23548	3.6

	409819	AK001015	Hs.55220	BCL2-associated athanogene 2	3.6
	430044	AA646450	Hs.152812	ESTs	3.6
	430829	AW451989	Hs.194024	ESTs	3.6
5	434224	AA360731	Hs.84	interleukin 2 receptor, gamma (severe co	3.6
	435247	AF080520	Hs.40767	EST	3.6
	431542	H83010	Hs.5740	ESTs	3.5
	430713	AA351647	Hs.2642	eukaryotic translation elongation factor	3.5
	434988	AA18055	Hs.161180	ESTs	3.5
	437748	AF234882	Hs.5814	suppression of tumorigenicity 7	3.5
10	418322	AA284186	Hs.84113	cyclin-dependent kinase inhibitor 3 (CDK	3.5
	439689	AW602168	Hs.222399	CEGP1 protein	3.6
	459583	AB07673		gbtL-BT152-080399-004 BT152 Homo sapiens	3.5
	403212			NM_019595.Homo sapiens Intersectin 2 (IT	3.5
	409099	AK000725	Hs.50579	hypothetical protein FLJ20718	3.5
15	453968	AA847843	Hs.62711	Homo sapiens, clone IMAGE:3351265, mRNA	3.5
	436338	W92147	Hs.118394	ESTs	3.5
	422890	Z43784	Hs.75983	ankyrin 3, node of Ranvier (ankyrin G)	3.5
	442295	AB27248	Hs.224398	Homo sapiens cDNA FLJ11469 fis, clone HE	3.5
	417575	AA641838	Hs.30085	hypothetical protein FLJ23186	3.5
20	433730	AK002135	Hs.3542	hypothetical protein FLJ11273	3.5
	438626	AW014875	Hs.137007	ESTs	3.5
	445873	AA250970	Hs.251945	poly(A)-binding protein, cytoplasmic 1-	3.4
	410153	BE311928	Hs.15830	hypothetical protein FLJ12891	3.4
	422128	AW881145		gbtCVO-OT0033-010400-182-e07 OT0033 Homo	3.4
25	414921	BE390551	Hs.77628	steroidogenic acute regulatory protein r	3.4
	441134	W28082	Hs.7878	cellular retinoic acid-binding protein 1	3.4
	444564	AI167877	Hs.143716	ESTs	3.4
	402470	NA		Target Exon	3.4
	418120	AA213437	Hs.192249	ESTs	3.4
30	422414	AW878237	Hs.13701	ESTs	3.4
	433345	AB181540	Hs.152982	hypothetical protein FLJ13117	3.4
	408213	U81412	Hs.51133	PTK6 protein tyrosine kinase 6	3.4
	422611	AA158177	Hs.118722	fucosyltransferase 8 (alpha 1,6) fucos	3.4
	423554	M30518	Hs.1674	glutamine-fructose-6-phosphate transamin	3.4
35	442619	AA447492	Hs.20183	ESTs, Weakly similar to AF184793.1 prote	3.4
	402359	NA		C19001991* <sup>g</sup> g128561111gbtAAK00751.1AF23.4	3.4
	438398	AA284267	Hs.221504	ESTs	3.4
	415208	P01020	Hs.172004	filin	3.4
	452853	AA812833	Hs.10845	ESTs	3.4
40	428345	R11141	Hs.198085	hypothetical protein	3.4
	449027	AJ271218	Hs.22880	dipeptidylpeptidase III	3.4
	412115	AK001763	Hs.73239	hypothetical protein FLJ10901	3.4
	432180	Y18418	Hs.272822	RuvB (E coli homolog)-like 1	3.4
	428977	AK001404	Hs.194688	cyclin B2	3.4
45	431611	U58786	Hs.264426	tissue specific transplantation antigen	3.4
	418286	AA622528	Hs.319825	Homo sapiens, clone IMAGE:3616574, mRNA	3.4
	435895	AF037538	Hs.5338	carbonic anhydrase XII (tumor antigen H	3.4
	443378	AF039250	Hs.9280	proteasome (prosome, macropain) subunit,	3.4
	428450	NM_014791	Hs.184339	KUAA0175 gene product	3.3
50	449571	AW016812	Hs.200266	ESTs	3.3
	412777	A135573	Hs.270123	ESTs	3.3
	420542	NM_005055	Hs.1321	coagulation factor XII (Hageman factor)	3.3
	412754	AW160375	Hs.74585	amyloid beta (A4) precursor-like protein	3.3
	418327	U70370	Hs.84136	paired-like homeodomain transcription fa	3.3
55	449065	AB272993	Hs.258983	ESTs, Weakly similar to high mobility gr	3.3
	425999	AW513051	Hs.332881	ESTs, Weakly similar to I38022 hypothet	3.3
	430280	AA361258	Hs.237868	Interleukin 7 receptor	3.3
	407777	AA161071	Hs.71465	squalene epoxidase	3.3
60	428516	BE262960	Hs.170187	glutamic-oxaloacetic transaminase 2, mit	3.3
	414361	AB06138	Hs.204044	ESTs	3.3
	427080	AW060287	Hs.173468	ras-related C3 botulinum toxin substrate	3.3
	426429	X73114	Hs.168349	myosin-binding protein C, slow-type	3.3
	448163	AA028880	Hs.25232	prolactin receptor	3.3
	428566	U41783	Hs.184916	calthrin, heavy polypeptide-like 1	3.3
65	418541	BE243138	Hs.86947	a disintegrin and metalloproteinase doma	3.3
	435283	AB01188	Hs.120910	ESTs	3.3
	411257	AA628887	Hs.115274	ESTs, Highly similar to IHH_HUMAN INDIAN	3.3

	430253	AK001514	Hs.236844	hypothetical protein FLJ10652	3.3
	430066	AI929659	Hs.237826	signal recognition particle 72kD	3.3
	436469	AK001465	Hs.5196	Down syndrome critical region gene 2	3.3
	437786	BE142681	Hs.155573	polymerase (DNA directed), eta	3.3
5	444079	H08046	Hs.23666	ESTs	3.3
	457163	H91082	Hs.118569	DY-binding protein IDAX (inhibition of	3.3
	431215	AA469078	Hs.121554	Human DNA sequence from clone RP11-216C13.3	3.3
	424563	AA46902	Hs.151426	ret finger protein 2	3.3
	450828	AW270955	Hs.193004	ESTs	3.3
10	408582	R43409	Hs.6829	Homo sapiens mRNA for KIAA1644 protein,	3.3
	445142	AW978484	Hs.93842	Homo sapiens cDNA: FLJ22554 fis, clone H	3.3
	425761	AI015709	Hs.172089	Homo sapiens mRNA; cDNA DKFZ56812022 (3.3	3.3
	439237	AW408156	Hs.316893	ESTs, Weakly similar to A47592 B-cell gr	3.3
	422616	BE300330	Hs.118726	selenophosphate synthetase 2	3.3
15	443247	BE614387	Hs.333893	c-Myc target JPO1	3.3
	406653	U24663	Hs.302063	immunoglobulin heavy constant mu	3.3
	434137	AA907734	Hs.124895	ESTs	3.3
	408877	AA479033	Hs.130315	ESTs, Weakly similar to A47592 B-cell gr	3.3
20	436101	C01765	Hs.36750	hypothetical protein FLJ11526	3.3
	406221	AA121683	Hs.47447	ESTs	3.3
	447619	U46255	Hs.339665	ESTs	3.3
	404755	NA		Target, Exon	3.2
	451871	AB21005	Hs.118599	ESTs	3.2
	420319	AW406289	Hs.96583	hypothetical protein	3.2
25	430580	AA086105	Hs.300697	immunoglobulin heavy constant gamma 3 (G	3.2
	400202	NA		NM_002795*Homo sapiens proteasome (pro3.2	3.2
	400222	NA		NM_002082*Homo sapiens G protein-couple	3.2
	425988	BE45897	Hs.274454	ESTs, Weakly similar to I38202 hypoteli	3.2
30	458096	BE550224	Hs.74170	metallothionein 1E (functional)	3.2
	430589	AJ02744	Hs.246315	UDP-N-acetyl-alpha-D-galactosamine:polyp	3.2
	431563	AD27843	Hs.120912	ESTs	3.2
	442353	BE379594	Hs.49136	ESTs, Moderately similar to ALU7_HUMAN	3.2
	422309	U79745	Hs.114524	solute carrier family 16 (monocarboxylic	3.2
	419703	AI793257	Hs.128151	ESTs	3.2
35	420380	AA409891	Hs.102408	ESTs	3.2
	410853	H04589	Hs.30469	ESTs	3.2
	454417	AI244459	Hs.110826	trinucleotide repeat containing G	3.2
	432745	AI821826		ghn17805.x5 NCI_CGAP_P03 Homo sapiens	3.2
	422032	AA476966	Hs.110857	polymerase (RNA) II (DNA directed) poly	3.2
40	415339	NM_015156	Hs.78396	KIAA0071 protein	3.2
	426384	AI472078	Hs.303662	ESTs	3.2
	448030	N30714	Hs.325060	membrane-spanning 4-domains, subfamily A	3.2
	418739	AA310964	Hs.88012	SHP2 interacting transmembrane adaptor	3.2
	442053	R35343	Hs.24968	Human DNA sequence from clone RP1-233G163.2	3.2
45	434747	AA637065	Hs.220585	ESTs	3.2
	427297	AW292593	Hs.334907	Homo sapiens, clone MGC:17333, mRNA, cont3.2	3.2
	412228	AW503785	Hs.73792	complement component (3d) Epstein Barr vl	3.2
	462304	AA025336	Hs.61511	ESTs, Weakly similar to SY559 cysteine	3.2
	453953	AW408337	Hs.36972	C27 antigen (p41)	3.2
50	407753	D50915	Hs.36365	KIAA0125 gene product	3.2
	451140	AL047586	Hs.10283	RNA binding motif protein 85	3.2
	430015	AW768399	Hs.112157	ESTs	3.2
	433313	W20128	Hs.298039	ESTs	3.2
	418334	AA319233	Hs.5521	ESTs	3.2
55	450223	AA418204	Hs.241493	natural killer-tumor recognition sequenc	3.2
	454365	AW966728	Hs.54642	methionine adenosyltransferase II, beta	3.2
	451128	AL116666		gbcDKFZp7810310_r1 761 (synonym: hary2)	3.2
	417793	AW405434	Hs.82575	small nuclear ribonucleoprotein polypept	3.2
60	428027	U22029	Hs.334345	cytochrome P450, subfamily IIA (phenobar	3.2
	441197	BE244638	Hs.166	sterol regulatory element binding trans	3.2
	424634	NM_003613	Hs.151407	cartilage intermediate layer protein, nu	3.2
	419395	AI345455	Hs.78915	G3-binding protein transcription factor,	3.2
	416714	AF283770	Hs.79630	CD74 antigen (immunoglobulin-associated	3.2
	449465	NM_004380	Hs.235958	CREB binding protein (Rubinstein-Taybi s	3.2
65	422186	W72424	Hs.112405	S100 calcium-binding protein A9 (calgran	3.2
	405079	W87707	Hs.82055	Interleukin 6 signal transducer (gp130,	3.2
	423551	AA327598	Hs.233785	ESTs	3.2

	453983	AA036849	Hs.61629	Homo sapiens cDNA FLJ12763 fis, clone NT	3.2
	442680	AT33682	Hs.130239	ESTs	3.2
	458079	AF19870	Hs.54277	DNA segment on chromosome X (unique)	9923.2
	425700	AF076292	Hs.169251	forlind box H1	3.2
5	417124	BC122762	Hs.25338	ESTs	3.2
	407104	BS7296	Hs.323810	v-erb-b2 avian erythroblast leukemia v	3.2
	442215	AI03172	Hs.129005	ESTs, Weakly similar to 2109260A B cell	3.1
	430271	T08199	Hs.237506	DnaJ (Hsp40) homolog, subfamily B, membe	3.1
	425317	AW205118	Hs.210546	Interleukin 21 receptor	3.1
10	426085	AI28023	Hs.89896	ESTs	3.1
	442313	BE388898	Hs.8216	hypothetical protein FLJ11307	3.1
	424709	AL137589	Hs.152149	hypothetical protein DKFZp434K0410	3.1
	429671	BE379336	Hs.211694	proteasome (prosome, macropain) 26S subu	3.1
	432715	AA247152	Hs.200483	ESTs, Weakly similar to KIAA1074 protein	3.1
15	431574	AW572699	Hs.261373	hypothetical protein dJ434O14.3	3.1
	436876	AI124756	Hs.5337	isocitrate dehydrogenase 2 (NADP), mito	3.1
	465017	NA		Target Exon	3.1
	433805	AI106910	Hs.112742	ESTs	3.1
20	437352	AL353967	Hs.284181	hypothetical protein DKFZp434P0531	3.1
	430105	X70297	Hs.2540	cholinergic receptor, nicotinic, alpha p	3.1
	422083	NM_001141	Hs.111266	arachidonate 15-lipoxygenase, second typ	3.1
	413507	BE145360	Hs.190064	ESTs, Weakly similar to 138022 hypotheti	3.1
	415989	AI267700	Hs.317684	ESTs	3.1
	422907	AI679283	Hs.6886	Human glucose transporter pseudogene	3.1
25	425548	AA890023	Hs.1908	prolactin receptor	3.1
	422569	BE387202	Hs.116638	non-metastatic cells 1, protein (NM23A)	3.1
	439963	AW247529	Hs.6793	platelet-activating factor acetylhydrola	3.1
	453883	AI638516	Hs.22630	cofactor required for Sp1 transcriptions	3.1
	458021	AI885190	Hs.156089	ESTs, Weakly similar to repressor protel	3.1
30	418478	U38945	Hs.1174	cyclin-dependent kinase inhibitor 2A (me	3.1
	403814	NA		Target Exon	3.1
	402327			Target Exon	3.1
	416635	AA190712	Hs.106875	gpcr3700r1 Stratagene HeLa cell s3 93	3.1
35	439836	AL355722	Hs.106875	Homo sapiens EST from clone 35214, full	3.1
	437036	AI571514	Hs.133022	ESTs	3.1
	449523	NM_000579	Hs.54443	chemokine (C-C motif) receptor 5	3.1
	406642	AJ245210	Hs.106875	gpcr3700r1 Stratagene HeLa cell s3 93	3.1
	406524	AF082762	Hs.106875	gpcr3700r1 Stratagene HeLa cell s3 93	3.1
	421924	BE514514	Hs.109606	coronin, actin-binding protein, 1A	3.1
40	414523	AU076633	Hs.76353	serine (or cysteine) proteinase inhibiti	3.1
	416379	N38867	Hs.203933	ESTs	3.1
	422823	D89674	Hs.121102	vanin 2	3.1
	433904	AI399956	Hs.208956	ESTs	3.1
	421904	BE143533	Hs.109309	hypothetical protein FLJ20035	3.1
45	428854	AW699713	Hs.339515	ESTs	3.1
	436043	AW693638	Hs.168630	Homo sapiens cDNA FLJ12136 fis, clone MA	3.1
	452823	AB011214	Hs.306956	transcription factor-like 5 (basic helix	3.1
	406381	NA		Target Exon	3.1
	428746	AW503820	Hs.192861	Sp1-B transcription factor (Sp1-1/P1.1 r	3.1
50	435147	AL133731	Hs.4774	Homo sapiens mRNA; cDNA DKFZp761C1712 (3.1	3.1
	425782	U68468	Hs.159525	cell growth regulatory with EF-hand doma	3.1
	423305	W88562	Hs.108198	ESTs	3.1
	419123	AA234276	Hs.88263	ESTs	3.1
	436581	AW977736	Hs.292133	ESTs, Moderately similar to I78885 serin	3.1
55	417105	X60992	Hs.81226	CD6 antigen	3.0
	428361	NM_015905	Hs.183658	transcriptional intermediary factor 1	3.0
	417880	BE241595	Hs.82848	selectin L (lymphocyte adhesion molecul	3.0
	402606			NM_004499*Homo sapiens hypothetical prot	3.0
	401451			NM_004499*Homo sapiens hypothetical prot	3.0
60	421878	AA299652	Hs.111496	Homo sapiens cDNA FLJ11643 fis, clone IE	3.0
	409518	BE384636	Hs.3454	KIAA1621 protein	3.0
	416933	BE561850	Hs.80506	small nuclear ribonucleoprotein polypept	3.0
	414324	Y14768	Hs.860	lymphotxin beta (TNF superfamily, membe	3.0
	425081	X74794	Hs.154443	minichromosome maintenance deficient (S	3.0
65	401519			C15000476-g[12]37279[af]XP_012163.1	3.0
	411704	AI499220	Hs.71573	hypothetical protein FLJ110074	3.0
	428819	AL135623	Hs.183914	KIAA0575 gene product	3.0

	428423	AU076517	Hs.184276	solute carrier family 9 (sodium/hydrogen	3.0
	413836	AI272727	Hs.249163	fatty acid hydroxylase	3.0
	412800	L28624	Hs.74101	spleen tyrosine kinase	3.0
5	410491	AA465131	Hs.84001	Homo sapiens clone 25218 mRNA sequence	3.0
	435858	L03078	Hs.156110	immunoglobulin kappa constant	3.0
	427685	A0791495	Hs.180142	calmodulin-like side protein	3.0
	452514	A0904898		gb:RC-BT068-130399-085 BT068 Homo sapie	3.0
	428500	X78585	Hs.289114	hexabrachion (tenascin C, cytotactin)	3.0
	432485	N90866	Hs.276770	CDW52 antigen (CAMPATH-1 antigen)	3.0
10	437400	AB011542	Hs.55589	EGF-like-domain, multiple 5	3.0
	482234	AW084176	Hs.223296	ESTs, Weakly similar to I38022 hypotheti	3.0
	413269	BE167526		gb:CM4-HIT0509-080300-107-g07 HIT0509 Homo	3.0
	453216	AL137566	Hs.32405	Homo sapiens mRNA; cDNA DKFZp586G0321 (f	3.0
	400929			ENSP00000252232*Storol regulatory elome	3.0
15	445145	AI661702	Hs.147434	ESTs	3.0
	432615	AA567191	Hs.55026	ESTs, Weakly similar to I54374 gene NF2	3.0
	423279	AW569861	Hs.290943	ESTs	3.0
	426392	AL109712	Hs.266506	Homo sapiens mRNA full length insert cDN	3.0
20	408548	AA055449	Hs.63187	ESTs, Weakly similar to ALUC_HUMAN III	3.0
	461345	R04006338	Rs.26312	glioma amplified on chromosome 1 protein	3.0
	413109	AW389945	Hs.110855	ESTs	3.0
	401714	NA		ENSP00000241802*CDNA FLJ11007 FIS, CLON	3.0
	421462	AF018495	Hs.104624	aquaporin 9	3.0
	421750	AK000768	Hs.107672	hypothetical protein FLJ20761	3.0
25	463293	AA382267	Hs.10683	ESTs	3.0
	457085	AA412446	Hs.98138	ESTs	3.0
	436930	AW843633	Hs.306163	hypothetical protein AL110115	3.0



TABLE 22A

**Table 22A** shows the accession numbers for those pkeys lacking unigeneID's for Table 22. For each probeset, we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Pkey: Unique Eos probeset Identifier number  
 CAT number: Gene cluster number  
 Accession: Genbank accession numbers

Pkey	CAT number	Accessions
407980	103087_1	AA046309 AI263500 AA046397
410789	1221055_1	AW803341 AW803265 AW803403 AW803466 AW803402 AW803413 AW803268 AW803396 AW803334 AW803355
411743	1250098_1	AW862214 AW869811 AW862215
412138	1279172_1	AW895367 AW895547 AW895564 AW896323 AW895405 AW895538 AW895538
413269	1366981_1	BE167526 BE167651 BE078401 R24654
416935	163179_1	AA190712 AA190605 AA252564
422128	211994_1	AW881145 AA490716 M85637 AA304575 T06067 AA331991
423945	233566_1	AA410943 AW848953 AA334202 AA332882
424109	235506_1	AW406678 AW966580 AW966151 AW966496 AA336174 AA336376 AA336637
424126	235728_1	AW866163 AA335983 AA336011 AA335688 AA335973
425331	250199_1	AW862128 AA355353 AA427363
426876	273265_1	BE069341 AW748403 AL044891 AJ806240 AA393080
432745	353673_1	AI821926 AA658826 AA664492 AA635129 AI791191
441153	51084_2	BE562826 BE378727
448212	755099_1	AI475858 AW969013
451126	859865_1	AL118669 D78823 AJ782176
462314	820172_1	AI904898 AJ904649 AJ904699
456207	165078_1	AA193460

TABLE 22B

Table 22B shows the genomic positioning for those pkeys lacking unigene ID's and accession numbers in Table 22. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique number corresponding to an Ecos processet  
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:489-493.  
 Strand: Indicates DNA strand from which exons were predicted.  
 Nt\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	Nt_position
400814	8569925	Minus	72840-72924,74761-74849
400829	7651921	Minus	122033-122241,123463-124028
401045	8117819	Plus	90044-90184,91111-91345
401451	8634068	Minus	119926-121272
401619	8649315	Plus	157315-157960
401646	7657839	Minus	34886-35133
401714	6715702	Plus	96484-96681
401666	8018106	Plus	73125-73623
402327	7656895	Minus	108675-108770,109801-109910
402359	9211204	Minus	40403-41961
402408	9796239	Minus	110326-110491
402470	9797107	Plus	195129-195776
402542	9801558	Minus	67076-67594
402576	9884928	Plus	66350-66496
402606	9909429	Minus	81747-82094
403011	6693597	Minus	3468-3623
403212	7630897	Minus	158037-158210
403329	8516120	Plus	96450-96598
403396	8783692	Minus	49323-49652
403486	9968528	Plus	2888-3001,3199-3532,3655-4117
404347	9638195	Plus	74493-74829
404560	8539738	Minus	240569-241589
404755	7795327	Minus	53729-53946
405017	8552084	Plus	35551-36080
405346	2914717	Minus	43310-43462
405381	8006920	Minus	7936-8054
405801	2524321	Plus	63469-63694
405850	8164935	Plus	13871-14110
406163	9929734	Minus	12802-13089
406348	9265985	Minus	71754-71944

**TABLE 23: 320 GENES DOWN-REGULATED IN BREAST CANCER  
COMPARED TO NORMAL ADULT BREAST**

- 5 **Table 23** shows 320 genes down-regulated in breast cancer compared to normal adult breast. These were selected as for Table 22, except that the numerator was set to the median value for 12 non-malignant breast specimens, the denominator was set to the median value amongst the 73 breast cancers, the 90<sup>th</sup> percentile value amongst the 12 non-malignant breast specimens was greater than or equal 80 units, and the ratio was greater than or equal to 4.0 (i.e. 4-fold down-regulated in tumor vs. normal breast).

	Key:	Unique Eos probset identifier number			
	ExAccn:	Exemplar Accession number, Genbank accession number			
15	UnigeneID:	Unigene number			
	Unigene Title:	Unigene gene title			
	R1:	Ratio of 50 <sup>th</sup> percentile normal body tissue to 75 <sup>th</sup> percentile tumor			
20	Key	ExAccn	UnigeneID	UnigeneTitle	Ratio
	428722	U76456	Hs.190787	tissue inhibitor of metalloproteinase 4	22.4
	428846	NM_000230	Hs.194236	loptin (murine obesity homolog)	17.4
	445263	H57646	Hs.42586	KIAA1500 protein	15.4
25	416905	T28459	Hs.89485	carbonic anhydrase IV	15.0
	407228	M25079	Hs.155376	hemoglobin, beta	14.5
	417511	AL049176	Hs.82223	chordin-like	14.5
	436285	AL133615	Hs.172572	hypothetical protein FLJ20093	14.3
	412442	AI983730	Hs.26530	serum deprivation response (phosphatidyl	13.6
30	410544	AA465433	Hs.95511	ESTs	12.6
	412047	AA934589	Hs.49696	ESTs	12.2
	422867	H25642	Hs.133471	ESTs	12.0
	406684	L34041	Hs.9739	glycerol-3-phosphate dehydrogenase 1 (so	12.0
	423201	NM_000183	Hs.126180	growth hormone receptor	11.7
35	422163	AF027208	Hs.112360	prominin (mouse)-like 1	10.8
	428769	AW207175	Hs.106771	ESTs	10.6
	407049	X72632	NM_021724*	Homo sapiens nuclear receptor	10.1
	412255	AW088626	Hs.117176	poly(A)-binding protein, nuclear 1	9.8
	425126	N32759	Hs.172944	chondric gonadotropin, beta polypeptide	9.8
40	406761	AJ220584	Hs.272572	hemoglobin, alpha 2	9.5
	447471	AF039843	Hs.18676	sprouty (Drosophila) homolog 2	9.5
	451533	NM_004657	Hs.26530	serum deprivation response (phosphatidyl	9.4
	419407	AW140377	Hs.41502	hypothetical protein FLJ21276	9.0
	411999	AI365585	Hs.145246	ESTs	9.0
45	410532	T30808	Hs.155376	hemoglobin, beta	8.9
	425707	AF115402	Hs.11713	E74-like factor 5 (ets domain transcript	8.8
	415585	X54162	Hs.79385	leiomyodin 1 (smooth muscle)	8.7
	443050	D78874	Hs.8944	procollagen C-endopeptidase enhancer 2	8.6
	435285	AA779958	Hs.185932	ESTs	8.5
50	422511	AU276442	Hs.117939	collagen, type XVII, alpha 1	8.4
	433138	AB024906	Hs.59729	semaphorin sem2	8.3
	402155		NM_004497*	Homo sapiens hepatocyte nuclei	8.1
	423350	A754634	Hs.131987	ESTs	8.1
55	445107	AJ208121	Hs.147313	ESTs, Weakly similar to I38022 hypotheti	8.1
	405843	N77976	Hs.272572	hemoglobin, alpha 2	8.0
	410199	AW377424	Hs.205126	Homo sapiens cDNA: FLJ22687, clone H	8.0
	417225	AA815048	Hs.24078	hypothetical protein FLJ12549	7.8
	437589	AA780849	Hs.294052	ESTs	7.5
	436082	AK000027	Hs.99633	ESTs	7.5
60	425078	NM_002599	Hs.154437	phosphodiesterase 2A, cGMP-stimulated	7.5
	430327	AW873636	Hs.55931	ESTs	7.4

	447577	A1336893	Hs.183297	DKFZP566F2124 protein	7.4
	446039	A1150491	Hs.50758	ESTs	7.2
	422060	R20693	Hs.326823	ESTs, Moderately similar to ALU5_HUMAN A	7.2
	424455	AA452006	Hs.333199	ESTs	7.1
5	424343	AW956360	Hs.4748	adenylate cyclase activating polypeptide	7.1
	442792	A1352340	Hs.131194	ESTs	7.0
	406714	A1219304	Hs.283108	hemoglobin, gamma G	6.9
	407571	A1448183	Hs.9572	ESTs, Highly similar to CYA5_HUMAN ADENY	6.8
	429580	AA346839	Hs.209100	DKFZP434C171 protein	6.7
10	453500	AI047427	Hs.43125	esophageal cancer related gene 4 protein	6.7
	422233	AB002058	Hs.113275	purinergic receptor P2X-like 1, orphan r	6.7
	420205	AA266395	Hs.88165	ESTs	6.6
	404368	NA		ENSP00000241079:TRRAP PROTEIN.	6.6
	447281	NM_008991	Hs.17917	extracellular link domain-containing 1	6.5
15	417090	AA193262	Hs.85863	ESTs, Weakly similar to B34612 zinc fing	6.5
	410677	NM_003278	Hs.65424	tetranectin (plasminogen-binding protein	6.5
	427263	AL119758	Hs.174185	adenonucleotide pyrophosphatase/phosphodi	6.5
	415011	AW963085		gb:EST375168 IMAGE reassesses, MAGH Hom6.4	6.4
	412088	S72043	Hs.73133	metallothionein 3 (growth inhibitory fac	6.4
20	416253	BE250659	Hs.15463	Homo sapiens, clone IMAGE:2958994, mRNA	6.4
	435885	AA701483	Hs.36341	ESTs	6.3
	402779	NA		Target Exon	6.3
	418138	AA213626	Hs.136204	EST	6.3
	438335	AA742687	Hs.62492	ESTs, Weakly similar to B39066 prolina-r	6.3
25	427019	AA001732	Hs.173233	hypothetical protein FLJ10970	6.2
	411478	BE143068		gb:MR0-1159-030200-003-b09 HT0158 Homo6.2	6.2
	452654	BE004783		gb:MR2-BN0114-270400-004-e11 BN0114 Homo6.1	6.1
	447359	NM_012093	Hs.18285	adenylate kinase 5	6.1
	414523	NM_014759	Hs.334658	KIAA0273 gene product	6.1
30	441266	H15968	Hs.293845	Homo sapiens, clone IMAGE:3502329, mRNA,	6.1
	417011	F08212	Hs.234658	ESTs, Weakly similar to Z109250A B cell	6.0
	400699	NA		Eos Control	6.0
	433614	W07475	Hs.277101	cytochrome c oxidase subunit IV isoform	5.9
35	440439	N52818	Hs.54754	ESTs, Weakly similar to potential CDS [H	5.9
	454404	BE067414		gb:MR4-BT0355-20010C-201-e05 BT0355 Homo.5.9	5.9
	438704	AA052610	Hs.148050	EST	5.9
	409563	NA		Target Exon	5.9
	433490	AW451023	Hs.65848	hypothetical protein DKFZp761O132	5.9
	419313	AA843367	Hs.87279	ESTs	5.9
40	409195	NM_001874	Hs.334873	carboxypeptidase M	5.8
	410882	AW809163		gb:MR4-ST0118-261099-012-e03 ST0118 Homo.5.8	5.8
	453469	AB014533	Hs.33010	KIAA0633 protein	5.8
	441899	A1372589	Hs.8022	TUGA protein	5.8
	426210	AA372052	Hs.334569	Homo sapiens cDNA FLJ14458 fis, clone HE	5.8
45	413055	BE063555		gb:CM1-BT0283-08189-033-b09 BT0283 Homo.5.8	5.8
	454102	AW873813	Hs.3343	phosphoglycerate dehydrogenase	5.7
	425187	AW014496	Hs.22509	ESTs	5.7
	429157	AW452355	Hs.258037	ESTs	5.7
	420202	AL036557	Hs.95910	putative lymphocyte G0/G1 switch gene	5.7
50	416284	AB68473	Hs.298006	ESTs	5.7
	428553	AA181641	Hs.184907	G protein-coupled receptor 1	5.6
	404689	NA		Target Exon	5.6
	438887	R68857	Hs.265499	ESTs	5.6
	406082	S47833	Hs.82927	adenosine monophosphate deaminase 2 (iso	5.6
55	449748	H23963	Hs.32043	ESTs	5.6
	431048	R50263	Hs.249129	cell death-inducing DFFA-like effector a	5.5
	452205	C15619		gb:C15619 Clontech human aorta polyA mRNA	5.5
	430400	AW446163	Hs.288609	hypothetical protein FLJ20159	5.5
	407744	AB020629	Hs.38095	ATP-binding cassette, sub-family A (ABC1	5.5
60	459606	A168905	Hs.60380	ESTs, Moderately similar to ALU6_HUMAN A	5.5
	414629	AA345824	Hs.76688	carboxylesterase 1 (monocyte/macrophage	5.5
	401655			C11000703.gi10048448[ref]NP_065268.1)g	5.5
	438107	T99079	Hs.191194	ESTs	5.5
	444432	A1161428	Hs.75916	splicing factor 3b, subunit 2, 145KD	5.5
65	434715	BE006346	Hs.116410	ESTs	5.5
	447205	BE617015	Hs.11006	ESTs, Moderately similar to T17372 plasm	5.5
	408122	A432652	Hs.42624	hypothetical protein FLJ10718	5.5

	454016	AW016806	His.233108	ESTs	5.5
	414913	R25621		gbyh4506.r1 Soares placenta Nb2HP Homo	5.4
	459033	AA017590	His.129907	ESTs	5.4
	441003	BE172240	His.126379	ESTs, Weakly similar to I39022 hypothesis	5.4
5	450637	N48626	His.18802	ESTs	5.4
	442388	AA994520		gbou42y09.s1 Soares_NF1_T_GBC_S1 Homo	5.4
	403612	NA		Target Exon	5.3
	407102	AA007629	His.9739	glyoxal-3-phosphate dehydrogenase 1 (so	5.3
	410067	R66634	His.268107	multimerin	5.3
10	428232	BE272452	His.183109	monoamine oxidase A	5.3
	432769	AA620814	His.144959	ESTs	5.3
	431344	R99530	His.272572	hemoglobin, alpha 2	5.3
	427032	AF012023	His.173274	integrin cytoplasmic domain-associated p	5.3
	406305	BE261320	His.158198	transcriptional adaptor 3 (ADA3, yeast h	5.3
15	437411	AW613948	His.194915	ESTs	5.3
	442600	A080481	His.131227	ESTs	5.3
	402054	NA		Target Exon	5.3
	432085	AF121829	His.272405	potassium channel, subfamily K, member 9	5.3
	416313	R59638	His.5181	ESTs	5.2
20	459159	A1904546		gb-QV-8T065-020339-103 BT065 Homo sapien	5.2
	427164	AB037721	His.173871	KIAA1300 protein	5.2
	441391	BE457930	His.170381	ESTs	5.2
	458959	A1285901	His.181297	ESTs	5.2
	402698	NA		ENSP00000251338*DJ1003J2.1 (sodium and	5.2
25	401810	NA		Target Exon	5.2
	438879	AA827674	His.189073	ESTs	5.2
	414557	AA24074	His.76780	protein phosphatase 1, regulatory [inhib	5.2
	427809	M26380	His.180878	lipoprotein lipase	5.1
	456353	NM_006744	His.76461	retinol-binding protein 4, interstitial	5.1
30	451186	AW023409	His.85285	ESTs, Weakly similar to leucine-rich gli	5.1
	451882	A0821324	His.100445	ESTs	5.1
	402583	NA		NM_021620:Homo sapiens PR domain contain	5.1
	431130	NM_006103	His.2719	HE4; WFDC2; putative ovarian carcinoma m	5.1
	458218	A1435179	His.126820	ESTs	5.1
35	416083	R53467	His.269122	ESTs, Weakly similar to ALU1_HUMAN ALU S	5.1
	456282	BE143867		gbLMR-HT0154-070100-013-h02 HT0164 Homo	5.1
	426488	X03350	His.4	alcohol dehydrogenase 1B (class I), beta	5.1
	426156	BE244537	His.167382	natrilin peptide receptor A/guanlylate	5.1
40	407891	AA486620	His.41135	endomucin-2	5.0
	408610	AW026692	His.224829	ESTs	5.0
	445967	D59597	His.118821	CGI-82 protein	5.0
	434813	A1524307	His.162870	ESTs	5.0
	437526	A076012	His.121388	ESTs, Weakly similar to MDHC_HUMAN MALAT	5.0
	454775	BE160229		gb-QV1-HT0413-090200-052-ar12 HT0413 Homo	5.0
45	409451	AF012626	His.54472	fragile X mental retardation 2	5.0
	408853	AW502327		gbUJ-HF-BR0p-aka-a-07-0-UJ.r1 NIHMGCC_5	5.0
	406062			Target Exon	5.0
	449490	AK000706	His.15125	hypothetical protein FLJ20699	5.0
50	417622	AW298163	His.82318	WAS protein family, member 3	5.0
	421978	AJ243662	His.110196	NICE-1 protein	5.0
	440338	R62431	His.12758	ESTs	5.0
	415421	R35009	His.24903	ESTs	5.0
	417574	R00348		gbye69e06.r1 Soares fetal liver spleen	5.0
	408882	AJ243191	His.56874	heat shock 27kD protein family, member 7	5.0
55	447988	A1768289	His.304389	ESTs	4.9
	445813	BE550889	His.158491	ESTs	4.9
	443074	AW341470	His.144507	ESTs	4.9
	451324	A1753000	His.208052	ESTs	4.9
60	432433	AW014734	His.157899	ESTs	4.9
	448554	A1989812	His.199850	ESTs	4.9
	414519	N94587	His.55083	ESTs	4.9
	457531	AW973716	His.13913	KIAA1577 protein	4.9
	433200	AA682722	His.192725	ESTs	4.9
	430782	AF026263	His.247920	cholinergic receptor, muscarinic 5	4.8
65	427555	AW137094	His.97990	ESTs	4.8
	433545	AA888510	His.112495	ESTs	4.8
	420334	A1349351	His.118944	hypothetical protein FLJ22477	4.8

	421795	X83904	Hs.283822	Rhesus blood group, D antigen	4.8
	427138	N76724	Hs.173717	phosphatidic acid phosphatase type 2B	4.8
	413072	BE083965		gb QV3-BT0296-140200-065-h1 BT0296 Homo	4.8
5	443721	AW450461	Hs.296355	ESTs	4.8
	408053	AW139474	Hs.246862	ESTs	4.8
	427057	AA843716	Hs.177927	ESTs	4.7
	442969	AI025498	Hs.132238	ESTs	4.7
	426220	AI383475	Hs.171697	ESTs, Weakly similar to T13924 sdk prote	4.7
	414593	BE386764		gb 601273249F1 NIH_MGC_20 Homo sapiens c	4.7
10	426893	AA396716	Hs.97418	ESTs	4.7
	434046	AW262618	Hs.113011	ESTs	4.7
	401590	NA		Target Exon	4.7
	457971	AW134679	Hs.242849	ESTs	4.7
	427722	AK000123	Hs.180479	hypothetical protein FLJ20116	4.6
15	443793	AA046290	Hs.26930	ESTs, Weakly similar to 2108230A B cell	4.6
	407737	RA8182	Hs.65959	ESTs	4.6
	441955	AA972327	Hs.142903	ESTs	4.6
	441499	AW298235	Hs.101689	ESTs	4.6
	447517	AI382726	Hs.182434	ESTs	4.6
20	403017			Target Exon	4.6
	450580	N40087	Hs.15248	ESTs	4.6
	404811	H56589	Hs.35155	Homo sapiens cDNA FLJ11027 fls, clone PL	4.5
	414831	M31158	Hs.77439	protein kinase, cAMP-dependent, regulate	4.6
	459290	NM_001545	Hs.34853	Inhibitor of DNA binding 4, dominant neg	4.6
25	444341	AI142027	Hs.146950	ESTs	4.6
	408614	AL137598	Hs.46531	Homo sapiens mRNA; cDNA DKFZp434C1915 [I4.8	4.6
	449538	AW204277	Hs.250723	hypothetical protein MGC2747	4.6
	434418	AF134707	Hs.278579	a disintegrin and metalloproteinase doma	4.6
	447350	AI375984	Hs.167216	ESTs	4.6
30	419583	F00312		gb H5880D101 STRATAGENE Human skeletal m4.8	4.6
	440598	AI349455	Hs.147492	Homo sapiens cDNA FLJ11777 fls, clone HE	4.6
	451169	AI290653	Hs.124750	ESTs	4.6
	438338	NM_014861	Hs.6168	KIAA0703 gene product	4.6
35	433756	AW015933	Hs.112654	Homo sapiens, clone MGC:9764, mRNA, comp	4.5
	423301	SE7580	Hs.1645	cytochrome P450, subfamily IVA, polypept	4.5
	417237	H86385	Hs.61737	palmitoyl-protein thioesterase 2	4.5
	439745	AL389981	Hs.149219	Homo sapiens mRNA full length insert cDN	4.5
	424137	AA335769	Hs.16252	ESTs	4.5
	449338	H73444	Hs.394	adrenomedullin	4.5
40	434744	N94835	Hs.283828	Homo sapiens genomic DNA, chromosome 21a	4.5
	407402	AF036303		gb Homo sapiens clone 23943 mRNA sequenc	4.5
	443510	NM_012190	Hs.9520	formyltetrahydrofolate dehydrogenase	4.5
	416754	AA169114	Hs.12247	hypothetical protein FLJ11413	4.5
45	416986	Z43619		gb HSC10E121 normalized Infant brain cDN	4.5
	457416	BE142062	Hs.62654	kringle-containing transmembrane protein	4.5
	418064	BE387287	Hs.83384	S100 calcium-binding protein, beta (neur	4.4
	437120	AI368125	Hs.157767	ESTs, Weakly similar to HXA2_HUMAN HOMEO	4.4
	453950	AA156996	Hs.211568	eukaryotic translation Initiation factor	4.4
50	401093			C12000566*gb 6330167 db BAAB5477.1 A	4.4
	436935	AW205494	Hs.253560	ESTs	4.4
	457974	AW842363	Hs.321717	ESTs, Weakly similar to S22765 heterogen	4.4
	428222	AL133112	Hs.183085	Homo sapiens mRNA; cDNA DKFZp434K098 [r	4.4
	442705	AE245434	Hs.131127	ESTs	4.4
	437409	AL359699	Hs.283860	Homo sapiens mRNA; cDNA DKFZp547C126 [r	4.4
55	458494	AB309508	Hs.158436	ESTs	4.4
	410480	H03589		gb yH2d08.r1 Soares placenta N2bHP Homo	4.4
	416059	R37101	Hs.20982	ESTs	4.4
	438463	AA807958	Hs.314232	ESTs	4.4
	444043	AI499723	Hs.139089	ESTs	4.4
60	438327	H67407	Hs.172944	chorionic gonadotropin, beta polypeptide	4.4
	457711	AF147401	Hs.23917	ESTs	4.3
	400870			C11000805*gb 11692665 gb AAG39879.1 AF28	4.3
	441425	AA833580	Hs.28937	homeobox protein from AL590526	4.3
	416287	H45384		gb yS9c10.r1 Soares adult brain N2bH5	4.3
65	407282	MI2873		gb Human Ig rearranged H-chain mRNA VDJ4	4.3
	444567	AV654020	Hs.184261	ESTs, Weakly similar to T26896 hypothet	4.3
	403263	NA		Target Exon	4.3

	410034	BE067414		gblcMR4-BT0355-200100-201-g05 BT0356 Homo	4.3
	456904	AI421645	His.139851	casein 2	4.3
	458427	BE395260	His.309438	EST	4.3
	416931	D45371	His.80485	adipose most abundant gene transcript 1	4.3
5	421296	NM_002836	His.103253	perilipin	4.3
	400973	NA		ENSP00000236657*-Muslin 5B (Fragment).	4.3
	452602	AW366194	His.55962	ESTs	4.3
	412330	NM_005100	His.788	A kinase (PRKA) anchor protein (gravin)	4.3
	405016			CY000171*-g[18280405]g[AA66402.1]AF245	4.3
10	436104	AI475671	His.88607	ESTs, Highly similar to F-box protein FB	4.3
	406118			ENSP00000246632:CDNA FLJ120261 fls, clone	4.3
	416558	T02850		gblcFB12A9 Fetal brain, Stratagene Homo s	4.3
	429745	AA480818	His.221736	ESTs	4.3
	433036	AW451206	His.115599	ESTs	4.3
15	444445	AA342229	His.115920	Homo sapiens cDNA: FLJ22816 fls, clone K	4.3
	453880	AI803196	His.28462	ESTs, Weakly similar to I38022 hypothet	4.3
	447384	AI377221	His.40528	ESTs	4.2
	414541	BE293116	His.76392	aldehyde dehydrogenase 1 family, member	4.2
	444975	AV652165	His.182462	ESTs, Weakly similar to T00352 hypothet	4.2
20	403621	NA		C5000212*-g[10047237]db[AB13407.1] (A	4.2
	451477	AI798425	His.42710	ESTs	4.2
	406344			C5001660:g[11811537]db[AB18935.1] (AB	4.2
	416970	AA191201	His.35851	DKFZP886E1021 protein	4.2
	413692	BE155868	His.26522	XIAA1803 protein	4.2
25	468504	AW070634	His.144764	ESTs	4.2
	404882	NA		C9001186*-g[12738842]pe[NP_037325.1] p	4.2
	418039	N99913	His.6858	ESTs, Weakly similar to I78655 serine/th	4.2
	404333			NM_001622:Homo sapiens alpha-2-HS-glycop	4.2
30	446532	AW975460	His.143563	ESTs	4.2
	414217	AI309298	His.279698	Homo sapiens cDNA: FLJ22165 fls, clone L	4.2
	418425	AI871247	His.6262	hypothetical protein MGC8407	4.2
	419589	AW937308	His.201925	Homo sapiens cDNA FLJ13446 fls, clone PL	4.2
	457029	AA397789	His.161803	ESTs	4.2
	447890	AF193807	His.131836	Rhesus blood group, B glycoprotein	4.2
35	448938	Y09763	His.22785	gamma-aminobutyric acid (GABA) A recepto	4.2
	440610	AI733068	His.130800	ESTs	4.2
	439590	AF086410		gblcHomo sapiens full length insert cDNA	4.2
	427240	AA395975	His.274151	Igulin	4.2
40	408932	AW594172	His.278513	TP53TG3 protein	4.2
	436112	T77545	His.187559	ESTs	4.2
	444382	AI144152	His.56246	ESTs	4.2
	459716	AA318060	His.135121	hypothetical protein FLJ22415	4.2
	415946	NM_015877	His.265581	Williams-Buren syndrome chromosome regl	4.1
	410039	R57171	His.57975	caldesmon 2 (cardiac muscle)	4.1
45	400545	NA		Target Exon	4.1
	403061	NA		Target Exon	4.1
	420139	NM_005357	His.95351	lipase, hormone-sensitive	4.1
	460244	AA007534	His.125062	ESTs	4.1
50	453261	AA034116	His.116494	ESTs	4.1
	440246	W62010	His.191379	ESTs	4.1
	414516	AI307802	His.135560	ESTs, Weakly similar to T43458 hypothet	4.1
	438232	AI150595	His.122226	ESTs	4.1
	410233	AA082847		gblcnn10g07.s1 Stratagene hNT neuron (937	4.1
55	412179	BE270758	His.69428	hypothetical protein MGC3020	4.1
	4141871	AK000705	His.153460	ESTs, Weakly similar to 1909123A Na gluc	4.1
	453392	AL110416	His.169764	hypothetical protein FLJ20701	4.1
	448640	AW817177	His.102558	gblcNF2p43AK9431_r1 434 (synonym: htes3)	4.1
	417481	AA032021	His.21768	Homo sapiens, clone MGC:5352, mRNA, comp	4.1
60	412912	AW118878	His.110835	ESTs	4.1
	454183	AW807116		gblcMR4-ST0062-040100-024-312 ST0062 Homo	4.1
	426328	AW631296		gblcHH83C09.y1 NCLCGAP_gu1 Homo sapiens	4.1
	435942	RG6285	His.191215	ESTs	4.1
	417629	T76945		gblcyc92c07.r1 Soares Infant brain 1N18 H	4.1
65	403593	NA		Target Exon	4.0
	402690			Target Exon	4.0
	418190	R49591	His.270425	ESTs	4.0

	408641	AW245207	Ha.5555	hypothetical protein MGC5347	4.0
	427099	AA629286	Ha.332053	serum amyloid A1	4.0
	446975	AB11536	Ha.145734	ESTs	4.0
	438831	BE263273	Ha.6439	synapsin II	4.0
5	455578	BE068350	Ha.14395	Homo sapiens cDNA FLJ13207 fls, clone NT	4.0
	401840	NA		Target Exon	4.0
	413753	U17760	Ha.75517	laminin, beta 3 (nicotin (125kD), kalinin	4.0
	445030	AI205925	Ha.147238	ESTs, Highly similar to AACS_HUMAN ALPHA	4.0
	433073	AW159913	Ha.150478	ESTs, Weakly similar to A Chain A, Cryst	4.0
10	456736	AW248217	Ha.1619	echaele-scute complex (Drosophila) homol	4.0
	459112	BE047734	Ha.5473	ESTs, Moderately similar to ALUS_HUMAN A	4.0
	448906	AI589567	Ha.308719	ESTs	4.0



TABLE 23A

Table 23A shows the accession numbers for those pkeys lacking unigeneID's for Table 23. For each probset, we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

Key: Unique Eos probset identifier number  
CAT number: Gene cluster number  
Accession: Genbank accession numbers

Key	CAT number	Accessions
409553	1156226_1	AW502327 AW502468 AW501629 AW502825 AW502687
410354	1170594_1	BE087414 BE087358 BE067419 BE067953 AW577127 AW501412
410233	118558_1	AA082647 AA083038
410480	1205347_1	H03588 AW750837 AW750888
410882	1225686_1	AW809163 AW808247 AW809177 AW809190 AW809225
411478	1247073_1	BE143088 AW849143 AW848705 AW848589 AW848071 AW848475 AW848092 AW848005
413065	1347960_1	BE083555 BE151321 BE151319 BE151657 BE151655 BE063556 BE151322
413072	1348163_1	BE063965 BE063988 BE064034 BE064026 BE063874 BE063956 BE063869 BE064043 BE064033 BE063884
414563	1484809_1	BE386764 BE387560
414913	1506721_1	R25621 C03959 CD4010
415011	151328_1	AW963085 AA159005 AW963073
415986	1504410_1	Z43619 RS1274 H12208 R12883
416287	1583547_1	H45384 H49125 H41699
417574	1687770_1	R03348 R08593
417629	1693382_1	T76945 R20210 R05755
418555	1767868_1	T02650
419583	186188_1	F00312 AA247490 F31427 AA383663 F22045
428328	284901_1	AW831298 AA375484
439590	47413_1	AF086410 W94386 W74609
442398	514127_1	AA994520 AW393574
452225	90415_1	C15819 AA024741 AA024742
452654	925831_1	BE004783 BE004847 AB11790
453692	977825_1	AL110418 AW676759
454183	1045936_1	AW807116 AW807569 AW807415 AW807338 AW807268 AW807263 AW807316 AW177402 AW807413 AW807068 AW141561 BE141559 AW807401 AW807310 BE141555 AW807315 AW807119 AW807299 AW807241 AW807225 AW807204 AW807345 AW807103 BE141615 AW807431 AW807393 AW807337 AW807406 AW807298 AW807375 AW845980 AW807220 AW807399 AW807064 AW807376 AW807024 BE141595 AW807236 AW807027 AW807112 AW807378 AW807202 BE141593 AW807216 AW807135 AW807244 AW807221 AW807297 AW807050 AW807249 AW807484 AW807075 AW807237 AW807212 AW807308 AW807110 AW807104 BE140912 AW807301 AW807382 AW807294 AW807025 AW807020 AW807108 AW807025 AW807433 AW807258 AW807201 BE141690 AW807302 AW807323 AW807380 AW807109 BE141588 AW845877 AW807418 AW807407 AW807309 BE141614 AW845881 AW807398 AW807300 AW807348 AW807311 AW807214 AW807132 AW807402 AW807350 AW807028 AW807298 AW807291 AW807305 AW807217 AW807312 AW807261 AW807266 AW807280 AW807266 AW807198 AW807346 AW807315 BE140940 AW807351 AW807122 AW807235 AW807076 AW807314 AW807213 AW807143 AW807131 AW800079 AW807086 AW807077 AW845823 AW807421 BE141619 BE140943 AW807420 AW807381 BE141621 AW845921 BE141629 BE141625 BE141624 BE141638 BE141630 AW907406 AW807290 AW807353 BE087414 BE067958 BE067419 BE067953 AW577127 AW501412 BE160229 AW8619879 AW8020179 AW819882 AW819876 AW820169 BE153201 AW993735 BE152911 BE143867 AW935060 AW880684 A1904646 BE179494 BE179421
454404	1170594_1	
454775	1234106_1	
455282	1273020_1	
458189	819898_1	

TABLE 23B

Table 23B shows the genomic positioning for those pkeys lacking unigene ID's and accession numbers in Table 23. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

10	Pkey:	Unique number corresponding to an Eos probeset		
	Ref:	Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham L et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham L et al., Nature (1999) 402:489-495.		
	Strand:	Indicates DNA strand from which exons were predicted.		
	NI_position:	Indicates nucleotide positions of predicted exons.		
15	Pkey	Ref	Strand	NI_position
	400545	9800107	Minus	124618-124981
	400870	9838306	Minus	34081-35027
20	400973	7890452	Minus	98119-98253
	401093	8516137	Minus	22335-23166
	401590	9998320	Minus	33547-33849
	401695	7145001	Plus	121591-122537
	401810	7342191	Plus	129063-129476
25	401840	7684597	Plus	56283-66439
	402054	8083691	Minus	8288-8805
	402195	7889778	Minus	147901-148884
	402553	7684486	Plus	94883-95003
	402650	8348058	Plus	13368-13598
30	402698	8570304	Minus	108641-108903
	402779	9988555	Minus	38173-39210
	403017	6693623	Plus	76630-79357
	403051	4827080	Minus	5269-5411
	403283	7770577	Plus	52431-52737
35	403433	9719811	Minus	72225-72437
	403593	6882850	Minus	62554-62712,69449-69602
	403612	8469080	Minus	94723-94859
	403921	7711590	Minus	3297-3535
	404358	7630958	Minus	102053-102199
40	404692	9797231	Minus	40977-41150
	404699	7534100	Plus	119461-119717
	405016	6524300	Plus	51997-53308
	405052	7857730	Plus	101253-101432
	405116	9143818	Plus	53597-54629
45	405344	8255974	Plus	20254-20374,20526-20859,20835-21097
	405593	7711604	Plus	34401-34538

TABLE 24:

Table 24 depicts Seq ID No., UnigeneID, UnigeneTitle, Pkey, Pred.Cell.Loc., and ExAccn for all of the sequences in Table 25. The information in Table 24 is linked by Seq ID No. to Table 25.

10	Pkey:	Unique Eos probe/ Identifier number			
	ExAccn:	Exemplar Accession number, Genbank accession number			
	UnigeneID:	Unigene number			
	Unigene Title:	Unigene gene title			
	Pred.Cell.Loc.:	Predicted Cellular Localization			
	Seq.ID.No.:	Sequence Identification Number found in Table 25			
15	Pkey	ExAccn	UnigeneID	Unigene Title	Pred.Cell.Loc. Seq. ID. No.
	449746	AI858594	Hs.176588	ESTs, Weakly similar to CP4Y_HUMAN CYTOD	Seq ID 1 & 2
	407276	AI951118	Hs.328738	Homo sapiens breast cancer antigen NY-BR	Seq ID 3 & 4
20	415639	AI733881	Hs.72472	BMP-R1B	Seq ID 5 & 6
	400297	AI127078	Hs.334473	hypothetical protein DKFZp56401278	Seq ID 7 & 8
	450375	AA009647	Hs.8850	a disintegrin and metalloproteinase domain	Seq ID 9 & 10
	102457	NM_001394Hs.2359		dual specificity phosphatase 4	Seq ID 11 & 12
	429170	NM_001394Hs.2359		dual specificity phosphatase 4	Seq ID 13 & 14
25	443399	AI905687	Hs.2533	aldehyde dehydrogenase 9 family, member	Seq ID 15 & 16
	422505	AL120802	Hs.124165	ESTs	Seq ID 17 & 18
	449765	NS2233	Hs.208332	ESTs, Moderately similar to ALUR_HUMAN A	Seq ID 19 & 20
	425692	D80041	Hs.155595	N-acetyltransferase 1 (arylamine N-acetyl)	Seq ID 21 & 22
	426215	AW663419	Hs.155223	staralpha2	Seq ID 23 & 24
30	439840	AW449211	Hs.105445	GNF family receptor alpha 1	Seq ID 25 & 26
	410102	AW248508	Hs.279727	Homo sapiens cDNA FLJ14035 fs, clone HE	Seq ID 27 & 28
	429220	AW207206	Hs.136319	ESTs	Seq ID 29 & 30
	416276	U41060	Hs.79136	LIV-1 protein, estrogen regulated	Seq ID 31 & 32
35	409079	W87707	Hs.82005	interleukin 6 signal transducer (gp130)	Seq ID 33 & 34
	442818	AK001741	Hs.8739	hypothetical protein FLJ10879	Seq ID 35 & 36
	442082	R41823	Hs.7413	ESTs	Seq ID 37 & 38
	444381	BE387335	Hs.283713	ESTs, Weakly similar to S84054 hypothetical	Seq ID 39 & 40
	446163	AA026880	Hs.25252	Homo sapiens cDNA FLJ13603 fs, clone PL	Seq ID 41 & 42
40	416636	NS2536	Hs.42845	solute carrier family 16 (monocarboxylate)	Seq ID 43 & 44
	433043	W57554	Hs.126919	ESTs	Seq ID 45 & 46
	429353	AL117406	Hs.200102	ATP-binding cassette transporter MRP8	Seq ID 47 & 48
	452190	H26735	Hs.91688	Homo sapiens clone PP1498 unknown mRNA	Seq ID 49 & 50
	446733	AA833600	Hs.28040	ESTs, Weakly similar to fatty acid omega	Seq ID 51 & 52
45	452747	BE153855	Hs.61460	Ig superfamily receptor LNIR	Seq ID 53 & 54
	423242	AL038402	Hs.125783	DEME-6 protein	Seq ID 55 & 56
	417433	BE270260	Hs.82126	5T4 oncofetal trophoblast glycoprotein	Seq ID 57 & 58
	432201	A1538613	Hs.296241	Transmembrane protease, serine 3	Seq ID 59 & 60
	423981	D13686	Hs.136348	osteoblast specific factor 2 (fascilin)	Seq ID 61 & 62
50	439869	AW602168	Hs.222399	CEBP1 protein	Seq ID 63 & 64
	114480	BE088776	Hs.151678	UDP-N-acetyl-alpha-D-galactosamine:polyp	Seq ID 65 & 66
	404581			NM_014112:Homo sapiens trichothrinophala	Seq ID 67 & 68
	328372	NA		Phase 2 & 3 Exons	Seq ID 69 & 70
	112267	AB033084	Hs.334806	KIAA1236 protein	Seq ID 71 & 72
55	355824	NA		ENSP00000249072:DJ22E13.1 (N-TERMINAL	Seq ID 73 & 74
	424735	U31875	Hs.272499	short-chain alcohol dehydrogenase family	Seq ID 75 & 76
	400289	X07820	Hs.2258	matrix metalloproteinase 10 (stromelysin)	Seq ID 77 & 78
	427585	D31152	Hs.179729	collagen, type X, alpha 1 (Schmid metaphase	Seq ID 79 & 80
	425925	NM_007695Hs.228213		cytochrome P450, 51 (lanosterol 14-alpha	Seq ID 81 & 82
60	428441	AJ24172	Hs.204098	lipophilin B (hemoglobin family member)	Seq ID 83 & 84
	421155	H67679	Hs.102267	lysyl oxidase	Seq ID 85 & 86
	420931	AF044197	Hs.100431	small inducible cytokine B subfamily (Cy	Seq ID 87 & 88
	420813	X51501	Hs.99949	prolactin-induced protein	Seq ID 89 & 90
	452744	AI267692	Hs.30504	Homo sapiens mRNA: cDNA DKFZp434E082 (fr	Seq ID 91 & 92

	420757	X78592	Hs.99915	androgen receptor (dihydrokolesterolone r	cytoplasm	Seq ID 93 & 94
	424905	NM_032487	Hs.153704	NIMA (never in mitosis gene a)-related k	nuclear	Seq ID 95 & 96
	429659	NM_007050	Hs.225952	protein tyrosine phosphatase, receptor t		Seq ID 97 & 98
5	446921	AB012113	Hs.16530	small inducible cytokine subfamily A (Cy	extracellular	Seq ID 99 & 100
	445537	AJ245671	Hs.12844	EGF-like domain, multiple 6		Seq ID 101 & 102
	428227	AA321649	Hs.2248	small inducible cytokine subfamily B (Cy	extracellular	Seq ID 103 & 104
	424001	W67883	Hs.137476	paternally expressed 10		Seq ID 105 & 106
	421727	Y13153	Hs.107318	kynurenine 3-monooxygenase (kynurenine 3		Seq ID 107 & 108
	452638	U65011	Hs.30743	preferentially expressed antigen in mela	nuclear	Seq ID 109 & 110
10	419867	AU077005	Hs.82208	a disintegrin and metalloproteinase doma		Seq ID 111 & 112
	414612	X72755	Hs.77367	monokine induced by gamma Interferon	extracellular	Seq ID 113 & 114
	426320	W47595	Hs.169300	transforming growth factor, beta 2	extracellular	Seq ID 115 & 116
	422967	L32137	Hs.1684	cartilage oligomeric matrix protein (pse	extracellular	Seq ID 117 & 118
15	411558	AA102670	Hs.70725	gamma-aminobutyric acid (GABA) A recepto		Seq ID 119 & 120
	417868	AW057903	Hs.82772	collagen, type XI, alpha 1		Seq ID 121 & 122
	426368	AJ249368	Hs.96558	ESTs		Seq ID 123 & 124
	431958	X53929	Hs.2677	cadherin 3, type 1, P-cadherin (placenta	plasma membrane	Seq ID 125 & 126
	428722	U76456	Hs.190787	tissue inhibitor of metalloproteinase 4		Seq ID 127 & 128
	412570	AB026436	Hs.177634	dual specificity phosphatase 10		Seq ID 129 & 130
20	421379	Y15221	Hs.103982	small inducible cytokine subfamily B (Cy	extracellular	Seq ID 131 & 132
	415752	BE314524	Hs.76776	putative transmembrane protein		Seq ID 133 & 134
	444951	N48373	Hs.10247	activated leukocyte cell adhesion molecu		Seq ID 135 & 136
	451110	AB95040	Hs.265398	ESTs, Weakly similar to transformation-r		Seq ID 137 & 138

**TABLE 24A**

- Table 24A** shows the accession numbers for those pkeys lacking unigeneID's for Table 24. For each probeset, we have listed the gene cluster number from which the oligonucleotides were designed. Gene clusters were compiled using sequences derived from Genbank ESTs and mRNAs. These sequences were clustered based on sequence similarity using Clustering and Alignment Tools (DoubleTwist, Oakland California). The Genbank accession numbers for sequences comprising each cluster are listed in the "Accession" column.

10    Pkey:            Unique Eos probeset identifier number  
      CAT number:    Gene cluster number  
      Accession:     Genbank accession numbers

15    Pkey    CAT number            Accession

335824 CH22\_3197FG\_619\_11\_LINK\_E  
 325372 c12\_hs

TABLE 24B

Table 24B shows the genomic positioning for those pkeys lacking unigene ID's and accession numbers in Table 24. For each predicted exon, we have listed the genomic sequence source used for prediction. Nucleotide locations of each predicted exon are also listed.

Pkey: Unique number corresponding to an Eas probedet  
 Ref: Sequence source. The 7 digit numbers in this column are Genbank Identifier (GI) numbers. "Dunham I. et al." refers to the publication entitled "The DNA sequence of human chromosome 22." Dunham I. et al., Nature (1999) 402:469-495.  
 Strand: Indicates DNA strand from which exons were predicted.  
 NL\_position: Indicates nucleotide positions of predicted exons.

Pkey	Ref	Strand	NL_position
404561	9795980	Minus	68036-70100

Table 25

The 69 gene sequences identified to be overexpressed in breast cancer may be used to identify ending regions from the public DNA databases (or used in GenBank). The sequences may be used to either identify genes that encode known proteins, or they may be used to predict the coding regions from genomic DNA using exon prediction algorithms, such as FGENISH (Salamov and Solovyev, 2000, Genome Res. 10:516-522)

Seq ID NO: 1 DNA sequence

Nucleic Acid Accession #: FGENISH predicted ORF

Coding sequence: 1-1518 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
1  ATGGAGGCGCT CCGGCTTCA GGAACCTCAT GCTCACCCCT TCTGCTGCT GATCCTGCT 60
2  TGACATGCTC TGCTGCTGTT TCAGGTAA TC AGGTTGTACC AGAGGAGGAG ATGGATGATC 120
3  AGAGGCGCTG ACGTGTGTTG TCACCCCTT GCGCCAGTGT TCTATGCCA CAAGGAGTT 180
4  TACCCAGTAA AGGAGTTTGA GGTGTATCAT AAGCTGATG AAAAAATCCC ATGTGCTGTT 240
5  CCGTTGTGGG TTGGACGCTT TAGCATGTTT TCAAGTGTC ATGACCCAGA CTAAGCCAA 300
6  ATTCCTCTGA AAGACAAGA TCCCAAAAGT GCTGTAGCC ACACAAATCT TGAATCCTG 360
7  TTGGTGCTGC GACTTGTGAC COTGGATGTT TCTAAATGGA AAAAGCACCC CCAGATTGTG 420
8  AAACTGTGCT TCAACATGAG CATTTCTGAA ATATTCATCA CCGATATGTC TGAGAGTGT 480
9  CGGATATGTC TGACAAATG GAGGGAAGCC ATTCGCCAAA ACTCACTCT CAGAGCTCTT 540
10 CAACATGTC CCGTAGTAC COTGGACAGC ATCATGAAGT GTGCTTCAG CGACGAGGCG 600
11 AGCATCCAGT TGGACAGTAC COTGGACTCA TACCTGAAAG CAGTGTTCAC CCITAGCAAA 660
12 ATCTCAAAAC AGCGCATGAA CAATTTTCTA CATCACAAGC ACCTGGTTT CAATTCAGC 720
13 TCTCAAGGCC AATCTTTTTC TAAATTTAAC CAAGAAGCTC ATCAGTTCAC AGAGAAGATC 780
14 ATCCAGAGCC GGAAGAGCT TCTTAAGAT AGCTTAAAC AAGATATGAT TACAGAAAGG 840
15 CGCTGGGATT TTCTGACAT ACCTTTGAGT GCCAAAGCC AAAACACCAA ACATTTCTCT 900
16 GAAAGCAGAT TCCAGGCTGA AGTGAAGAGC TTCAATGTTT CAGGACATGA CACCACATCC 960
17 ATGCTATCT CCGTAGCTCT TTAATGCTTG GCAAAGTACC CTGAGCATCA CGAGAGATCG 1020
18 CAGATGAGAA TCAGGAACTC CTGAGGGAT GGGTCTTCTA TTACTTGGA ACACCTGAGC 1080
19 CAGATGCTCT ACACACGAT GTGCTATCAG GAAATGCTCC GCGCTAGG ACCGATGTA 1140
20 AACATATCCC GGTACTCGA CAAGACCACT ACCTTTCCAG ATGGAGCGCT CTATCTTCA 1200
21 GGAATAACTG TGTTATCAAA TATTGGGCT CTTCACCACA ACCGCTATT CTGGGAAGAC 1260
22 CCTCAGGTC TTAACCCCTT GAGATCTCC AGGGAATAAT CTGAAAAAT ACATCCTAT 1320
23 GCGTTCATAC CATTCACAG TGGATTAAAG AACTGCTATT GGCAGCATTT TGCCATAATT 1380
24 GAGGTAAAG TGGCAGTGG ATAACTCTG CTCCAGCTCA AGCTGCTCC AGACATCA 1440
25 AGGCTCCC AGGCTGTGTC TGAAGTGTTC TCAAGTCA AGAATGGAT CCATGTGTT 1500
26 GCAAAAAAAG TTTCCTAAT TTAAGTCTT TCGTATAAGA ATTAATGAGA CAATTTCTCT 1560
27 ACCAAAGGAA GAACAAGAG ATAAATATAA TACAAAATAT ATGTATATGG TTGTTTGACA 1620
28 AATTATATAA CTAGGATATC TCTGACTGG TTTCGACAT CATTAACAGT AATTTTAATT 1680
29 TCTTTGTTAT ACTTGTTGAA ACCCAAAAAA ACACCTGUAA AAACCTCAAGC TGACTTCGAC 1740
30 TGCAAAAGGA AATTATGTT TTGTGTAAT AGTGATGAG TTGGCTTCAA GCATAGCTGT 1800
31 ATCAAACCT CACTAGTAT CTGCATTAAT TTATCTCTG CAAATATCTG CATGATAGCT 1860
32 TTATTCTGAG TTATCTTCC CCAATAATAA AAAA

```

Seq ID NO: 2 Protein sequence

Protein Accession #: FGENISH predicted

```

1      11      21      31      41      51
1  MEPSWLRLRM AHPELLLLL CMSLLLPQV KLYORRRWMI RALHLFPAP AHVHFYGHKEF 60
2  YVVEEVEVYL KLMGKTPCAV PLVVGPFPMF FSVHDDPYAK ILKXKQDPKS AVSHKLELWS 120
3  VGRGLVTLDG SKWKKHQIV KPGENISLK IFITMMSESV RMMLNKWEER LAQNSRLIELF 180
4  QHVSMLTLDL IMKCAFESHQ SIQLDSTLDS YLKAVPNLSK ISNQRMNLF LKHNDLVKFS 240
5  SQGQIFSKFN QELHQFTKCV IQDRKESLKD KLKQDTTKR KWDPLDLS AKSHNDKFS 300
6  EADLQARVKT FMAHGHDTHS SAJWSLVCY AKYPBHQBC RDEIRELLOD GSNITVHSL 360
7  QMPYTIMCKR ECLLYAPTV NISLLDKPT TPDGNSLSA GIVTINWIA LHSNPVWED 420
8  PVDNVLPSB RENSEKIEY AFIFESAGLR NCGQHPAH ECKVAVALTL LRFKLAPDHS 480
9  RFPQVQRQVV LKSNQIHFV AKKVC

```

Seq ID NO: 3 DNA sequence

Nucleic Acid Accession #: NM\_052997

Coding sequence: 100-4125 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
1  CTATGCTATA CCAGCAACGA CTCCTACATC GTCCACTCTG GGGATCTTGA AAGATGCAAT 60
2  AAAGCTGCTT CCGGGGGA CA AGTCCGGAAG CTGGAGAAGA TGACAAGAG GAAGAAGACC 120
3  ATCAACCTTA ATATACAAGA CGCCCAAGA AGGACTGCTC TACACTGGCG CTGTGTTCAAT 180
4  GGCCATGAGG AAGTAGTAA CATTCTGGTA GACAGAAAGT GCCAGCTTGA CTTCTTGAAT 240
5  GCGGACACCA GACACACTCT GATGAAGGCT CTACAAATGC ATCAGAGGCG TTGTGCAAT 300
6  ATTCTGATAG ATTCTGTTAT GATATAAAT CTCTGATGAT TGATAGGCA CATGCGCTC 360
7  CATTATCTCT TTATAGAGTGA GATTTTCTCA GTGTGGGAAA AACTGCTCTC CCATGTGCA 420
8  GTCATGGAAG TGCAACAAA GCGTAGCCTC ACACCACTT TACTATCCAT ACACAAGA 480

```

AGTGAACAAA TTGTGGAATT TTGTGATATA AAAAATGCCAA ATGCGAATGC AGTTAATAAG 540  
 TATAAATGCA CAGCCCTCAT GCTTGCTGTG TGTCTAGGAT CATCAGAGAT AGTTGGCGATC 600  
 CTCTCTCAGC AAAATGTGGA GCTCTTGTCT CGAGATATAT GTGGAATAC TCGAGAACAT 660  
 5 TATGCTGTTA CTGTGGGATT TCATCAGATT CATGAACAAA TTATGGAATA TATACGAAAA 720  
 TTAATCAAAA ATCATCAAAA TACCATAATCA GAAGGAACAT CTGAGGAAC ACTGTATGAG 780  
 GTGTCACCCCT TGGCGGAAAG AACACTGTAC ACAGCTGTAA GCTGTGTGGA AAAGAACACT 840  
 GATGTCGTGG CACCTTGCTT GGAAGAAGACA CCGTACACGG CTGAAGAAGCT GGTGGAAAAA 900  
 AACCTGTATG AGGCTGTATC CTGTGTGGAG GGAAACATCTG ACAAAATTCA ATGTTTGGAG 960  
 10 AAAGCGACAT CTGGAAGAAT GGAACATGCA GCAGAAGAAA CACCTAGGGA AATTACGAGT 1020  
 CTTGCAAAAAG AAACATCTGA GAAATTATCG TGGCCAGACAA AAGGAAGACAT TAGGAAGATC 1080  
 GCATGTGGAGA AAAAGAAGAA CACACCTAAG GAAATTATTA GTCCCGCAAA AGAACACATC 1140  
 GAGAAATTTA CTTGACAGAC AAAGAAGAGA CTAAGGAAGA TCGATGGGA GAAAAAAGAA 1200  
 ACACCTGTAA AGACTGTGAT CGTGGCAAGA GTAACATCTA ATAAACTAA ATGTTTGGAA 1260  
 AAGGAAGAT CTAAAGATAT TGCATGTCTC ACAAAAGAAAT CATCTACAAA AGCAAGTGCC 1320  
 15 AATGATCAGA GOTTCCCATC AGAATCCAAA CAGAGGAGAG ATGAAGAATA TTCTGTGAT 1380  
 TCTCGGAGTC TCTTTGAGAG TCTGCAAAAG ATTCAGTGTT GTATACCTGA GTCTATATAT 1440  
 CAAAAAGTAA TGAAGATAAA TAGAGAGTGA GAAGAGCCTC CTAGAGAGCC ATCTGCTCTC 1500  
 AAGCTTGCCA TTGAATGCA AAACCTGTGT CCAATAAAG CTTTGAAT GAAGAATGAA 1560  
 CAACAATGGA GAGCAGATCC GATGTGTCCA CCGAAGATCCA AACAAGAAGA CTATGAAGAA 1620  
 20 AATCTTGGG ATTTGAGAG ATCTGTGAG ACTGTTTTCAC AGAAGGATGT GTTGTTAACC 1680  
 AAGGCTACAC ATCAAAAAGAA AATAGATAAA ATAAATGGAA AATTGAAGAA GTTCTCTAAT 1740  
 AAAGAATGTC TTCTGAAGGC TACCTGTGGA ATGAAGAATT CTATTCCATC TAAAGCCTTA 1800  
 GAATTAAGG ACATGTAAC TTCTAAAGCT GAGCTCTCG GGAAGGCATC TGCTTTGAGC 1860  
 CTTGCACTGT AAATGCAAAA GTCTGTCCA AATAAAGCCT TGAATGTGAA AATGAGACAA 1920  
 25 ACATGGAGAG CAGATGAGAT ACTCCCATCA GAATCCAAAC AAAAGGACTA TGAAGAAGAT 1980  
 TCTTGGGATA CTAAGATCTT CTGTGAGACT GTTCCACAGA AGGATGTGTG TTACCCCAAG 2040  
 GTGTCGCATC AAAAGAAGAA AGATAAATAA AATGGAATAAT TAGAAGGCTC TCTGTATAA 2100  
 GATGTCTCTC TGAAGCTTAA TAGAGAGTGA TGTAAAGTTA TCTACTGA AAAGCCTTGA 2160  
 30 TGAATGACA TGCAAACTTT CAAGCAGAG CCCTCCGAGA AGCCATCTGC CTTGCAAGCT 2220  
 GCGATTGAAA TGCAAAAGTC TTTTCCAAAT AAAGCCTTGG AATTGAAGAA TGAAGAACAA 2280  
 TTGAGAGCAG ATGAATATCT CCCATCAGAA TCCAACACAA AGGACTATGA AGAAGTTCT 2340  
 TGTGATCTTG AGAGTCTCTG TGAGACTGTG TCAGAGTAAG ATGTGTGTTT ACCCAAGCT 2400  
 ACACATCAAA AGAATAAGTA TAAATTAAT GGAATAATG AAGAGTCTCT TGATATGAT 2460  
 35 GTTCTTCTGA AGCTGCTCTG CAGATGTAAA GTTCTATTTC CACTAAAGC CTTGAGATCT 2520  
 ATGACATGCA AAATCTTCAA AGCAGAGGCT CCCAGAGAAC CATCTGCTT CAGAGCTGCC 2580  
 ATTGAATGTC AAAAGTCTGT TCCAATAAAA GCTTGGAAAT TGAAGATGA ACAACATGAT 2640  
 AGAGCAGATC AGATTTGCC TCTGAGATCA AAACAAGAGA AGTTGTGAAG AAAATTCTGG 2700  
 GATTTGAGAG GTCTCGTGG GACTGTGTTC CAGAAGAGAT TGTGTGTGAC CAGGCTGACA 2760  
 CATGAAAAG AATGAGTGA AATAATGAGA AATTGAGAG ATTCACATAG CCTATCAAAA 2820  
 ATTTGTGATA CAGTTCATCT TTGTGAAAGA GCAAGGGAAC TTCAAAAGA TCACTGTGAA 2880  
 CACGCTACAG GAAAAATGGA ACAAAATGAA AAGAAGTTT GTGTACTGAA AAGAAGACT 2940  
 40 TCGAAGCAAA AAGAATAAAA ATCAGAGTTA GAGAACCAAA AAGTTAAATG GGAACAAGAG 3000  
 CTCTGACAGT TGAGATGATC TTTAAACCAA GAGAGAAGAG AGAGAAGAA TGCCGATATA 3060  
 TTAATGAAA AAATTAGGGA AGAATTAGGA AGATATGAG AGCCACATG GAAGAGATTA 3120  
 45 GAATGTAAAC AACAACTTGA ACAGGCTCTC AGAATACAG ATATAGAAT GTAGAGTGTA 3180  
 GAAAGTAAAT TGAATCAGGT TTCTCACACT CATGAAAATG AAAATTATCT CTATCATGAA 3240  
 AATTGATCT TGAAAAAGGA AATTGCCATG CTAAAACTGG AATAGGCCAC ACTGAAACAC 3300  
 CAATACCGAG AAAAGGAAAA TAAATACTTT GAGGACATTA AGATTTTAAA AAAAAAGAT 3360  
 50 GTGTAACTTC AGATGAGCCTT AAACTGTGAA GAGGAAGATC TAACTAAAG GGCATCTGAA 3420  
 TATAGTGGGC AGCTTAATCT TGTGTGTAGCT GAGAGACAAA TGCTCACTT TAAATGGA 3480  
 GAAAAACAG ACAGAAAGAT ACTAGAGCCA GAAATTGAAT CACACATCC TAGACTTGCT 3540  
 TGTGCTGTAC AAGACATGTA TCAAAATTGT ACATCAAGAA AAAGTCAAGA AACTGTCTT 3600  
 CACATGTGAG GAGATGTGCT TTGCAAGAA AAAATGAAT TTGATGTGAG TAGTAGATA 3660  
 55 TATAACAATG AGGTGTGCTA TCAACACTT TCTGAAGCTC AAAGGAATCT CAAAGCCTTA 3720  
 AAAATTAATC TCAATTATCG AGAGAGATCT CTAGAGAGAA ATCATGTGTG TTGAGACAT 3780  
 GACAAAGAG ACCAAGCTGA AACACAGTGT CAAATGAAG AGCTGTACA CATGTATCAA 3840  
 AACGAACAG ATAAATGTGA CAACACACAT GAACAGCAGG AGCTGTGTA TCAGAAATTA 3900  
 TTTCACACTC AAAGCAAAAA TATGTGGCTC CAACAGCAT ATAGTCTATG ACATAAGAAA 3960  
 60 GTGACACAA AAAGCAAGAT ACAATATGAT ATTCATTTC TTGAGAGGAA AATGCAACAT 4020  
 CATCTCTCTA AAGAGAAAAA TGAAGAGATA TTTAATTACA ATACACCTAT AAAAAGCCT 4080  
 ATATATAT ATGAATAAGA GAAGAAGACA AGCAATGCTT CATGAGAGC AAGCAGATAA 4140  
 AAATCTCTTT TTGAGAAAGA ACAGACCAGA TCTTTACTCA CACTCATCTG TAGGAGAGT 4200  
 65 GTCTAGCAT CACCTTATGT TGAATAATCT ACCAATAGCT TGTGTCAAC GAATACTTAT 4260  
 TTAGAAGAA AAATCATGTA TTTCTTCTG AAGCTACAG ACATAAATA ACAGTTGAA 4320  
 GAATTAAGT TTACAGAAAT GCATAAGCT CCACAGGATT CCACCTAGC CTGATGATC 4380  
 AGCAGACATC ATCAATCCA ACCAGAAATC CACTGTGAC TCAGGCTAG GTGACAGAT 4440  
 GAGACTCCAC CTGAGAAA

Seq ID NO: 4 Protein sequence  
 Protein Accession #: NP\_443723.1

1 11 21 31 41 51  
 75 MTKRRKKTDL NQDAQKRLTA LHWACVNGHIE EYVTVLVDKQ COLDLVDGEH RPTAKALQC 60  
 HQEACANILI DSVADNLYV YCYGNALHYA KLLBQIGAVY VIKKASLPL 120  
 LLSIKRSEQ IVEFLIKNA NANANYKYKC TALMLAVCHG SSEIVMLQL QNVDFVADI 180



COVTAEHYAV TCGPHHDEQ DMEYIRKLSK NQONTNPEGT SAGTPDEAAP LAERTYDTAE 240  
 SLVECTPDEA APVATRETFD AELVEKTFD EASLSVSTIS DKQCCLBKAT SGKFFQSAIE 300  
 THRUITSFAK FTSBKFTYPA KGPSPKIAWE KKBOTPREIM SPAKETSSEKF TWAAKGRFRK 360  
 IAWKKHTPV KTGCVARVTS NKTKVLEKGR SKMACPTICE SSTKASANDQ RFPSESKQBE 420  
 DEEYSDCSRS LFESSAKIQV CIPESIQKV MENNEVEEP PKKPSAFKA IEMQNSVPNK 480  
 APELKNEQTL RADPMPPFES KQKDYBNSW DSESLCETVS QKDYVCLPKAT HQKEIDKING 540  
 KLEBSPNKDG LLKATCGMKV SIPTKALELK DMQTKAEPP GKPSAEPEAT ENQKSVPNKA 600  
 LELKNBTWFL ADHIFESSE QKDYBNSWD TSELCTVRQ KDVCLPKAAH QKQEDKINGK 660  
 LBGSVPKDLG LKANCGMKVS IPTKALELMD MQTKFAEPFS KPSAFEPAE MQKSVFNKAL 720  
 ELKNEQTLRA DEILPESBQK KDYBESSWDS ESLCTVSQK DVCLPKATHQ KEIDKINGKL 780  
 EESPNDNDGFL KAPCRMVKS IPTKALELMDM QTKFAEPPEK PSAPFAPEIM QKSVPNKALE 840  
 LKNEQTLRAD QMFPEBSEKQK VBEENSWDS SLRETVSQKD VCVPKATHQK EMDKSGKLE 900  
 TWISLSKLD TVSEKSEHNS LQKHCEBQT GKMBWKKKT CVLKJKLSKA KEIKSLQNE 960  
 KVKWEQBLCE VRLTLNQEBE KERNADILNE KIRERLGRIE EQHRELEVK QQLLEQLRIQ 1020  
 DIELKSVESN LNQVSIHTEN ENYLLHENCN LKKBIAMLKL SIATLKHQYQ EKENKYFEDI 1080  
 KILKEKNABL QMTLKLKEES LTKRASQVSY QLKVLIAENT MLTSKLKBEQ DKEILEABIE 1140  
 SHHPRLASAV QDHDQVTSR KSQEPAPHIA GDACLQRMKN VDWSTYNNN EVLIHQLSSEA 1200  
 QKRSKSLCN LNYAGDALRE NTLVSEHAQR DQRELEQMHF EABHAYQNEQ DNVNKHQET 1260  
 EELDKQLFOL QSKNMAAQO LVHAHKEADN KSKTHDEHF LERKMQHLL KEKNEHFNY 1320  
 NNHLKNRIYQ YEKEKAETEN S

Seq ID NO: 5 DNA sequence

none found

Nucleic Acid Accession #:

Coding sequence: 275-1785 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CCGCGGCGCC GGAGTCGGCG GGGCGTCGCG GGAACCGCGC AGTGGCGAGA CCGCGGCGCT 60  
 GAGGACGCGG GAGCGGGGAG CGACGCGGGG GGGTGGAGTT CAGCCTACTC TTCTTGATG 120  
 GTGAAGAAGAA AGGAAGATCA TTTCATGCCCT TGTTGATAAA GGTTCAAGACT CTGCTGATG 180  
 CATAACCAAT TTGGCTCTGAG CTATGACAAG AGAGAGAAACA AAAAGTTAAA CTTACAAGCC 240  
 TGCCATAAGT GAGAAGCAAA CTCCTCTGAT AACATGCTTT TCGGAAGTGC AGGAAAATTA 300  
 LAATGTGGCA CCAAGAAGAA GATGTGTGAG AGTACAGCCC CCACCCCGCG TCCAAAGGTC 360  
 TTGCTGTGTA AATGCAACCA CCATTGTCCA GAGAGACTAG TCAACAATAA TTGACGACAA 420  
 GACGAGATAT GTTTCACGAT GATAGAAGAG GATGACTCTG GGTGTCCTGT GGTCACTTCT 480  
 GGTTCCTCTG GACTAGAAGG CTCAGATTTT CAGTGTGGG ACACCTCCAT TCCTCATCAA 540  
 AGAAGACCAA TTGATAGCTG CACAGAAGG AAGAGATGTA ATAAAGACTC ACACCTCTCA 600  
 CTGCTCTGAT TGAAAACAGC AGATTTGTGT GATGACCTA TACACAGAGC GCGTTTACTT 660  
 ATATCTGTA CTCTCTGAT TTGCTCTGT GCTCTATCA TATTAATCT TTACTCTCGT 720  
 TATAAAGAC AAGAACCAG ACCTGATAC AGCATTGGGT TAGAACAAGTA TGAACCTTAC 780  
 ATCTCTCTG GAGATACCTC GAGAGACTTA ATTGAGCAGT CTCAGAGCTC AGGAAGTGG 840  
 TCAGGCTCTC CTCTCTGGT CCAAGGACTT ATAGCTAAGC AGATTACAGT GGTGAAGAGC 900  
 ATTGAAAAAG GTCCGTATTG GGAAGTTTGG ATGGGAAAGT GCGCTGGCGA AAAGTATGCT 960  
 GTGAAAGTGT TGTTCACAC AGAGAGAAC ACCTGTGTCA GAGAGACAGA AATATATCAG 1020  
 ACAGTGTGA TGAAGCATTG AACAATTTTG GTTTTCATGT CTGCAATAT CAAAGGACAA 1080  
 GCGTCTTGGA CCGAGTTGTA CCTAATCACA GACTATCATG AAAATGGTTC CATTATGAT 1140  
 TATCTGAAGT CCACACACCT AGACGCTAAA TCAATGCTGA AGTTAGCTTA CTCTTCTGT 1200  
 AGGTGCTTAT GTCATTTACA CACAGAAATC TTATGACTC AAGGCAACCC AGCAATTGCG 1260  
 CAYCGAGATC TGAAAAGTAA AATCACTTGT GTGAGAAAAA ATGGAACCTG CTGATATGCT 1320  
 GACCTGGCGT TGGCTCTGA GTGAGACAGT GATGATGCTC TATGACAGAT CGGAAAGATC 1380  
 ACCTGAGTGT CCGCAACCAO CTATATGCT CCAGAAGTGT TGGACGAGG CTTGAACAAA 1440  
 AATCACTTCC AGTCTTACAT CATGGCTGAC ATGTATAGTT TTGGCTCAT CTTTGTGGAG 1500  
 GTGCTAGGA GATGTGTATC AGGAGGTATA GTGGAAGAT AACAGCTTCC TTATCATGAC 1560  
 TCACTGCCCA GTGACCCCTC TTATGAGGAC ATGAGGAGAA TTGTGTGCTT CAAAGAGTTA 1620  
 CCCCCTCAT TCCCAAAAGT GTGAGACAGT GATGATGCTC TATGACAGAT CGGAAAGATC 1680  
 ATGACAGAT GTCTGGCTCA CAATCTCTCA TCAAGAGCTGA CAGCCCTGCG GGTTAAGAAA 1740  
 ACACCTTGCCA AAATGTGAGA GTCCCAAGGAC ATTAACCTCT GATAGGAGAG GAAAAGTAAG 1800  
 CAATCTTGCA GAAAGGCCAC AGGTACTCTT CTGTTGTGG GCAAGAGCAA AAGACATCAA 1860  
 TAAGCATCCA CAGTACAAGC CTTGACATC GTCTCTGCT CCAGTGTGGT CAGACATCC 1920  
 CTTTCAGGGA GCGAGATGGC CAAAGACAGA GAAGTCCCA GAGAGAGAGA TTGATCCGTG 1980  
 TCTGTTGTA GCGGAGAAAA CCGTGGGTA ACTTGTCAA GATATGATGC AT

Seq ID NO: 6 Protein sequence:

Protein Accession #: none found

1 11 21 31 41 51  
 MLLRSAGKLN VGTKKEDGES TAPTRPKVL RCKCHHCPH DSVNNICSTD GYCFMIBED 60  
 DSGLFVYVSG CLGLGGSGGQ GDTTPPQKQ RSEICTERN ECKNQLHPTL PPKRNRDFVD 120  
 GPHERKALL SVYSCSELLY LILCPYRY KRQETRRYS IGLEQDETLY IPFESIRDLL 180  
 BQSSGSGSGS GLPLVQRTY IKQJQMVKY QKGRYGVWYM KWRGRBKVA VPFTEEBAS 240  
 WPRETEIYQT VLMRHENILG FIADIKGWT SWTQVLYLTD YHENGSLDYD LKSTLDAKS 300  
 MLKLAYSVS GLCHLHTEIF STGKQPAJAH RDLKSKNLV KKGNGCCAD LGLAVKFSID 360  
 TNEYDIPNPT RVGTKRYMFP SVLDESLNRN HFQSYMIADM YSGLFLVWV ARRCVSGGV 420

HEYLQPYHDL VPSDFSDYM REIVCIKKLR PFPNNRWSSD ECLRQMOKLM TECHWHPAS 480  
RLTALRVKKT LAKMSESQDI KL

Seq ID NO: 7 DNA sequence

Nucleic Acid Accession #: none found

Coding sequence: 482-3097 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
| | | | |  
10 AACTGAGCTA ACAAGAAATA CTAGAAAAAG AGGAAGGAGA ACATTCTGCT AGCTTGGATC 60  
TAAAGACTAA GAAGAGAGGA GTGATCAATG TCAGCTCTGT TAAACACTGT GTTTACTTAC 120  
TGCAATCAGC AGCTTGCAAA TGGTTAACTA TATGCAAAAA AGTCAGCATG AGCTTGAAGT 180  
ATGCGGTGAA TTTTAATTGA GGGAAAAAGG GACAATTGCT TCAGATGCTC CTAGTATGCA 240  
CTCTGCTTGA AATATTTTCA ATGAAATGCT CAGTATTCTA TCTTTGACCA GAGGTTTAA 300  
CTTTATGAAG CTATGGGACT TGACAAAAAG TGATATTTGA GAAGAAAGTA GCGAGTGGTT 360  
GGTGTGTCT TTTTATAAT AAGAGAAATG AATTATGTGT AACACCTCTT CCAGCTGTGC 420  
ATTACAGATA AGCTCAAGAA GAGTCTGTGC TTTACAGAA TGGAAITTCAT CACATGACAA 480  
CA1GAAAGTC TGGATTCAAT TC1TTTATTC ATCTCTCCIT GCTGTATAT CTTTACATC 540  
CCAAACTCCA GTGCTCTCAT CCAGAGGGCTC TTTGTATTCT CTTTGCATTT GTGAGGAAAA 600  
AGATGTGCAC ATGCTAATAA AT1GTGGAAG AAAAGGTATC AAGATGGTAT CTGAAATAAG 660  
TGTGTCACCA TCAGAGACCT TCACCAATAG CTTAATAAAT AACCGCTTGA CGATGCTTCA 720  
TCAGAAAGAC TTTCTGTGGC TTACCAATGC TATTGCAATA CACCTTGGAT TTACCAATAT 780  
TCAGATATT GAGATAGGTG CATTAAATGG CCTTGGCCCT CTGAAACAAC TTCATATCAA 840  
25 TCACAAATCT TTAGAATTC TTAAGAGGGA TACTTTCCAT GGACTGGAAG ACCTGGAAAT 900  
CTCGCAAGCA GATAACAATT TTATCACAGT GATTGAACCA AGTGGCTTGA GCAAGCTCAA 960  
CAGACTCAAA GTGTATAATT TAAATGACAA TGTATTGAG AGTCTTCTCT CAACAATCT 1020  
CCGATTTGCT CAGTAAAGCC ATCTAGATCT TGGTGGAAAT CAATTACAAA CATTCCTTAA 1080  
TTTGTGTTT CTGCAACACA TTGGCCGAAT ATTGGATCTT CAGTGGGAGG ACAAACAATG 1140  
30 GGCCTCGAAT TGTGACTTAT TGCAGTTAAA AACTTGGTTG GAGAACAATG CTCACAGCTC 1200  
TATAATTGGT GATGTGTCT GCAACAGCCC TCCATTTTAA TAAAGGAAGA TACTCAGTAG 1260  
ACTAAGAAAG GAATCTATT GTCCCTACTCC ACCAGTGTAT GAAGAACATG AGGATCTCTT 1320  
AGGATCATTG CATGCTGCAAT CAGCACTTTC AATAAATGAT A GTGGCATGT CACTAAGAC 1380  
CAGCTGCAAT CTAAAGACTAC GACCAAAAGC ACCAGGTGTT ATAGCTTATA TTCAAGAGCC 1440  
35 ATCCACTCAA CTTCCAGGAC CTTACTGGCC TATTCTGTGT AACTGCAAGG TCTTATCCCC 1500  
ATCAGGCAAT CTAATACAT GTACAGGAGCG CAACATTGAA AGCTTATCAG ATCTGAGAGC 1560  
TCTCCGCAAA AATCTAGAAA AGCTCATTCT AGCGGGAAT ATTTATCAAC GTTATAGAAA 1620  
GTCTGATCTA GTGGAATAT TTACCTTTGA AATGCTTLCAT TTGGGAACAA ATGATATTGA 1680  
AGTTTGTGTA GATGATGAT GTGATGATCT AAGGAGATTA CAAGAATCTT ATCTAAATG 1740  
40 TAAACACTCG ACCAAATTAA GTAAAGGACT CTGTCTGTGT CTTCAATACT TTGATATCT 1800  
ATATCTTGA TACAATGCC TTAAGGAAAT ACTGCCAGGA ACCTTTAATC CAATGCCATA 1860  
ACTTAAAGTC CTGTATTAAA ATAACAACCT CCTCCAGTGT TTAOCACCA ATATTTTTCT 1920  
AAGGTTTCTT CTAACATAAG TAAATCTTAA ACAAACCAAG TTTACCATC TACGTGTAAG 1980  
TAAATATTGG GATGATCTGT ATTACTAAC CCAGATGTGC CTTGAGGATA ACCCTGGGA 2040  
45 CTGCTCCTGT GACCTGTGTT GACTGTGAGA ATGGATACAA AAGTTAAGCA AGAACAGCT 2100  
GCAGAGTAC ATCTCTGCA CTTCCGCCGG GCATCTGCGA AAAAAGGAAT TGAAGGCCCT 2160  
AAATAGTGAA ATCTCTGTC CAGOTTATGT AAATAACCCA TCCATGCCAA CACAGACTAG 2220  
TTACTTATGT GTCCAGACTC CTGCAACAGC ACAAATAG GCTGATATTA TTTTACGATC 2280  
TCTTACGAG CTTGTGACAC TGTCTGTCT AATATTGGA CTTCTGATTA TGTTCATCAC 2340  
50 TATTGTTTC TGTGCTGAG GGAATGTTGT TCTTGTGTT CAGCCAGAGA GAGATACAAA 2400  
AAGAACAACA GTAGTAGAGC AAATGAGAGA CAACAGTCTT GTGCATCTTC AGTACAGCAT 2460  
GTATGCGCAT AAAACCACTC ATCACACTAC TGAAGAACC TCTGCTCAC TCTATGAACA 2520  
GCACATGTTG AGGCCATGTT TCAATGTCTA TAGAAGTCCA TCTTTGCTC CAAAGCATCT 2580  
55 GGAAGAGGAA GAAGAGAGGA ATGAGAAAGA AGGAAGTGT GCAAACATC TCCAAAGAG 2640  
TCTTTTGA CAAGGAAAAT ATGACCACTC CACAGGTGCA AATATGAAT ACAAACCA 2700  
GAACCAATCA ACAGAATTT TATCTTCCA AGATGCCAGC TCAATTGACA GAAACATT 2760  
AGAAAAGAAA AGGGAACCTC AGCAACTGGG AATCACAGAA TACCTAAGGA AAAACATG 2820  
TCAGCTCCAG CTTGATATGT AGGCACATTA TCTTGGAGCC CAGGAAGAGC TGAAGTAA 2880  
GGAACACATTA ATGTACTCAC GTCCAAAGGA GTTATTAGTG GAACAGACAA AAAATGAGTA 2940  
60 TTTTGAAGT AAAGCTAAT TACATCTGGA ACCATGACTT TTAGAAAGTCC TGGAGCAGA 3000  
ACATATAGTG GAGATATTGA GCGCTTTCGC AGGAATGCTG TGA TCTGTTT TTAAGTCCAT 3060  
ACCTTGTA AA TAAGTGCTT AGCTGAGTGT GTCATCAATC AGAACCTAAG CACAGACGTA 3120  
GACTATGGGG AAAAAAAG AAGAAGAAA GAAACTCAGG GATCACTGGG AGAAGCGATG 3180  
65 GAATTATCT CAGCGAATTT AGTCTGTCCC AAATAAATAT AATCTTGCA TGTAAATC

Seq ID NO: 8 Protein sequence:

Protein Accession #: none found

1 11 21 31 41 51  
| | | | |  
70 MKLWIHIFYS SLLACISLHS QTVLSSRRGS CDSLNCNEEK DGTMLNCEA KGKIMVSESS 60  
VPPSRFPQLS LLNNLGLTML TNDPGLTNA ISHLGFNNI ADIEHGFNG IGLKQLHIN 120  
HNSLEILKED TPGHLENLSP LQADNNFYT IEPSPSKLN RLKVLINDN AIESLPNIF 180  
75 RFPVLTHLDL RQNGQLTPLY VGLEBHGRJ LDLELDNWK ACNEDLIQLK TWLENNMPPS 240  
JIGDVCNSP RFGKSLSR LKESICPTT PYVBEHEDP GSLIHAAATSS INDSMSMTK 300  
TSILKLPYKA PGLPIYTRK STQLGPPYCP IPONCKVLP SGLLIHQER NIESLDRP 360

PPQNRKLLJL AGNIHSLMK SOLVEYFTLE MLHLGNHJUE VLEHSGMHL TRLQKLYLNG 420  
 NMLKLSKGM FLGKJL NLYLFPNPKL LPTFPNPKL LKVLNNNL LQYLPHFIS 440  
 5 GVLPLXVNLK TNQTHLPVS NILDDLILL QULDEPNWD CSDLVGLQK WQKLSLPHS 540  
 TDDILCTSPG HLDKKELKAL NSEILCPGLV NNPSPMTQTS YLMVYTPATT TNDADILRS 600  
 10 LIDAVPLSVL ILGLLIMPTI IVFCAAGIVV VLVHLRRRYK KEQVDEQMRD NSPVHLQYSM 660  
 YGHKTTHPTI EBPASLYIQ HMVSPMVHVV RSPSPGPKHL EBEERNEBK GSDAKHLQRS 720  
 15 ILBQNHSLPT TOSNMKYKIT NQSTEPJLHQ DASSLYRNL EKERLEQQLQ ITEYLARKNA 780  
 QLQPDMEAHY PGRHEELKLM ETLAKYSRPK VLVRETKNEY FELKANLHAE POYLEVLEQQ 840  
 T

Seq ID NO: 9 DNA sequence

Nucleic Acid Accession #: NM\_005474

Coding sequence: 307-3936 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CACTACGCT CTCTCTAGT CCGCGGCCAA CTGGACAGT TTGCTCATTT ATTGCAACGG 60  
 TCAAGGCTGG CTGTGTGCCA AACGGGGGCG GCGGACGCA CCGACACACA CGGGGGGAAA 120  
 20 CTTTTTAAA AATGAAGAAG TAGAAGAGCT CAGCGGCGCG GCGGGCCGTG CCGGAGGGCT 180  
 CCGGAGCTGA CTCGCCGAGG CAGGAAATCC CTCCGGTGCG CAGCCGCCGC CCGCGTGGC 240  
 GCGCGGTGGG GATGGGTGCG CGCTGCGCG GCGGCGCGAG AGCTGCTGCA CTGAAGGCCG 300  
 GCGCAAGAGG CAGCGCGCCG GCTGCGCGTG TCCCGCGCG GCGGCTCTCT CTGAGCGCT 360  
 25 GCGCGTCTCT CTGCGCGCG CTGCGAGCC CAGCGGTGTA GCTTATGAAA GAAAGGAAGA 420  
 GCTGTGAAG TTGTGACGTG CTCTGTTCGG AGTGCGGACC TCTGGATCCC AGTGAAGAG 480  
 TTGACTCCA AGAATCATCC AGAAGTCTCT AATATTGAC TACAACGGGA AAGCAAGAA 540  
 CTGATCATAA ATGTGGAAAG AATGAAAGGT CTCATTGACA GCAAGTTTAC GGAACCCACC 600  
 TATCTGCAAG ACGTACTACIA TGTCTCCCTC GTCTGAAA TT ACACGGTAAT TCTGGTTCAC 660  
 30 TGTTACTACC ATGGACATAT ACGGGATATT TGTATTGAG CAGTCAATCT CAGCAGCTGT 720  
 TTGCTGTCTA GGGGACTTAT TTGTTTGA AATGAAGACT ATGTCTTGA GAAACATGAA 780  
 AGTGTCAACA ACAGATACAA ACTCTTCCA GCGAAGAAGC TGAAGAAGGT CCGGGATCA 840  
 TTGTGTATCA ATCAACAAC ACCAAACCTC GCTGCAAGA ATGTGTTTCC ACCACCTCT 900  
 CAGCATATGG CAAGAAGCA TAAAGAGAG ACCTCAAGG CAACTAAGTA TTGTGAGCTG 960  
 35 GTAGTGTGG CAGACAACCG AGATTTTCAG AGGCAAGGAA AAGATCTTCA AAGGTATGAG 1020  
 CAGGCAATGA TAGAGATTGC TAATCAAGTT GACAAGTTT ACAGACCAT GACATATGG 1080  
 ATGTGTTTGG TAGGCTGTGA AGTGTGGAAT GACATGGACA AATGCTCTGT AAGTCAAGG 1140  
 CCATTACCA GCGCTCATGA ATTTCTGGAC TGGAGGAAGA TGAAGCTCT ACTCGCAAA 1200  
 TCCCATGACA ATCGCGAGCT TTGTAGTGGG GTTATTTC AAGGGACCA CATTGGCATG 1260  
 40 GCGCAATCA TGAAGTATG CACGCGACAG AGATGTGGG GAAATGTGAT GACCAATCA 1320  
 GACATATCC TTGCTGTGCG CCGTGAAGAT GGCATGAGC TGCGGACAAA TTTCGGGATG 1380  
 AATCATGACA CACTGTGACA GGGCTGTAGC TGTCAAATGT CGGTTGAGAA AGGAGCTGC 1440  
 ATCATGAACG CTTCACCGG GTACCAATT CCATGTGTGT TCAGCAGTTG CAGCAGGAAG 1500  
 GACTTGGAGA CCAGCGTGG AAGAGGAATG GGGGTGTGCG TGTTAACTC CGCGGAAGCT 1560  
 45 AGGAGTCTT TCGGGGCCA GAAATGTGG ACAGATTGT TGAAGAAGAG AGAGAATGT 1620  
 GACTGTGGG ACCCAGAGA ATGTATGAT GCGTCTGCA ATGCCAACG CTGTACCTG 1680  
 AARGCGAGCG CTGTGTGCG ACATGGGCTG TCTGTGAAG ACTCGACGT GAAGCTCTCA 1740  
 GGAAAGAGGT GCGAGGACT CAGCAACTCC TGTGACCTC CAGAGTCTG CACAGGGGCG 1800  
 AGGCTCACT GCCAGCCAA CGTGTACTG CAGCATGGC ACTCATGTA GGATGTGGAC 1860  
 50 GCGTACTGT ACAATGGCAT CTGCGAGACT CACGAGCAG AGTGTGTAC ACTCTGGGA 1920  
 CCAGGTGTCA AACTGTCCC TGGGATCTG TTGAGAGAG TCAATCTCG AGGTGATCT 1980  
 TATGCACT GTGCAATG CTGAAAGAT TCTTTGCA AATGGAGAT GAGAGATCT 2040  
 AATATGAAA AAGCAATG TCAAGGAGT GCCAGGCCG CAGTCAATG TACCAATGG 2100  
 GTTCCATAG AAACAACAT CCCCTGACG CAAGAGGCC GGATTCCTG CCGGGGAGC 2160  
 55 CAATGTACT TGGGCGATG CATGCGGAG CAGCGGCTG TGTCTGAGG CACAAGATGT 2220  
 GCAGATGAAA AATCTGCTT GAATCTGCA GTGCAAAATA TTATGTCTT TGGGTGTAC 2280  
 GACTGTGCA TGCATGCA CCGCAGAGG GTGTGCAACA CAGGAAGA CTGCCACTCG 2340  
 GAGGCGCACT GGGCACCTC CTCTGTGAC AAGTTTGGT TTGGAGGAAG CACAGACAG 2400  
 GGCCCATCC GGCAGACAGA TAACCAAGT TTAACCATG GAAATCTGT GACCATCTG 2460  
 60 TTCTTCTTG CTGCGGATT TTGTGTTAT CTCAAAAGAA AGACCTTAT ACGACTCTG 2520  
 TTACAAATA AGAAGACAC CATTGAAAA TTAAGTTGT TGCCTCTCT CGGCGCAC 2580  
 CTGCGCTCT AACCTGCGG AACAACCTGG CTACTCATG CTTTGGCGG GACCCGAC 2640  
 CCGCAGAT CTCTACCTG GAGAGACAT CCGAGAGAT TCTCTGAGT TCAAGATGT 2700  
 GACATCAGA GACCCCTCAA CGGCTGAA TTGCTTCAG CCCAGTCA CACAGCAGT 2760  
 CTCTCTCCC TCCACCGGG CCGACGTGCA CTAAGGCTC CTGCCAGCC CCGGCCAG 2820  
 65 AAGCTGAC CATTGCGAG CCGAGGGAGC TGAAGCCAA ACCGCCCTCA GAGCGCTCTG 2880  
 CTGCGAGTC CTGCGCGCA AACAACCTGG CTACTCATG CTTTGGCGG GACCCGAC 2940  
 CATGTGAGA CATTGCGAG CTGCGACCC CTGACAGCT CTCACAAATA TCCACAAA 3000  
 GTGCCAGAT CCACCCAC CCGCTATATT AATGAGAG CCGACACCT TTTTCAACG 3060  
 TGAAGACAGA AGTTTGCACT ATCTTTCAGC TCCAGTTGGA GTTTTGTGA CCACTTTTA 3120  
 70 GGATTTT TTAATGTTAA AACATCAAT CATAGAAGC TTGAGCTAC TGCGTCACT 3180  
 GGTGTGCTGT GCTATGTGTC TGTGTTACT TGCACAGTA CTGTAAAT ATTAATTA 3240  
 GCGAATATTT GATACAGTGT CAGTGGKXK TAGTAGCAT TTTACATC ACTGATTT 3300  
 CCATGGCAGG AAGGCTTGT GTGCTTTAG TATTTATGT AACCTGAAAT ATCTGTCTG 3360  
 75 ATGGGATCT GGACAGGATG TGTGTCCTT CTGATCAAG CCTATTGA AAGCATGCC 3420  
 CCAACTACCC CCAGCTGTGC TTATGTTACC AGATGCACT CAAGAGATCC CAAGTAGAT 3480  
 CTCAGTTGAT TTCTTGAT TTCCATCTCA GGCACAGACC AAGGGCTCT AGGTCCAGG 3540  
 TTGTTTGGC TTTCAGGAG GCCTGTGCG CTTGACAA TGGCAGGAG GCTCCAGG 3600

ACACCTTGGGA GAAATCTGGC TTCTGGCCAG GAAGCTTTTG TGAGAACCTG GGTTCGACAC 3650  
 AGGAACTCTTA AGGTGTAGCC ACACAGGATG AGAGAGCTGGA ACATGAGACA AGCGAGAAGT 3720  
 TGACCCCTUAG CTGACGAGCC GTGAGCATGT TTGGAAGGGG TCTGTATGT CTACTACGAG 3780  
 GGCTGCTTGT AGAAATGCCA AGCACTCTTT TTCTCGCTGT TCCTTCTAG AGCACTGCCA 3840  
 5 CCAATGAGGTT ATTATGCTTG GGAAAGGTGG TGTTCTCTGA AGAAACCTAC TGCCAGGACA 3900  
 CTGCAAAACCG CCACCTTCCCT ATACTGCTTG GAGCTGAGCA AATCACCAACA AACTGTAAAT 3960  
 CAATGATCTCT GTATGTAGAG AGATGAGGAC TTCTCATGCG ACCAACAATA TTCTTCAGAT 4020  
 TGAACACATTA ACCAGATCTA GTCAATCAAG TCTGTTTACT GCAAGGTCTA ACTTATTAAC 4080  
 AATTAGGCGAC ACCTCTTATG CTTCGCAAAA CTACAACCAA TGGAAATGTGA TGTTCTATGG 4140  
 10 TATAGTTCAT GTCTGCTATC ATTATTCGTA GATATTGGAC AAAGAAGCTT CTCTATGGGG 4200  
 CATCTCTCTT TTCAACACTTG GCTGCGAGGAA TCTTTAAAGG ATGCTTTTAA CAGAGTGTGA 4260  
 ACCCTATTTCT TAAACACTTG CAACCTACTT GTTGAGCATC ACAGAAATGTG ATAGAAGAAAT 4320  
 CAACCTGCTT ATCAACTGCT TAAATATTAT GAGATGTGGC TTGGCGACGA TCCTCCTGAA 4380  
 CTCTTCACCT TTCAATGGCC TGACTAGGGA GCCATGTTTC ACAAGGTCTT TAAAGTGACT 4440  
 15 AATGGCATGA GAAATACAAA AATACTCAGA TAAAGTAAAA TGCCATGATG CTTCTGTCTT 4500  
 CTGGACTGGT TTTCACATTA GAAGACAATG GACAACAGCT ACATAATTCA CTCTGATGTT 4560  
 TTATAGAGAA AGCCTCTCTT TGGGTCAAC AGTTTTCCTA TCGTCTGAAA CAGAAAATAA 4620  
 TCTACCAAGA ATCTGTGTTT GCTCTCAGA AACAAGAACT GCAATTCTACT TTCCGCTGT 4680  
 TCCCTACTGT ATCTAGGCCA CATAGTATTC ATGACTATGG ATAAACTATA CAGGTGACAC 4740  
 20 AAACACACAC AAAAGGGAAC CCAGCTCTAA TACATTCCAA CTGCTATAGC ATGCATCTGT 4800  
 TTATCTATA CTGTATTAAGT TCTTTAAAT GTAAAGGCCA CTGTGGAAAAT AATACCTGTG 4860  
 AGATACATAC AGAATTACTG TAAGTGATTA CAGCTGTGTA TGTACTATAA GCCAAACATA 4920  
 TATATACTAT TAAAGAGCT CACAGAACTT TATAGGCACT TACGTGGCCA TTCTCTTTT 4980  
 25 AGATGCCCAA ATCTCTAGT CTGCAATGTT AGCCCTTCTT CCAATTATAA GAGGATATGA 5040  
 ACCAAAAAAA AAAAAAATAA AA

Seq ID NO: 10 Protein sequence  
 Protein Accession #: NP\_033465.2

1 11 21 31 41 51  
 MAARLPVSP ARALLLALAG ALLAPCEARG VSIWNIGRAD EVVSASVHSG DLWVFKSFD 60  
 SKNPEVLNI RIQRREKLI INLEHRELLI ASPTETHTYL QDQTDVSLAR NYVILGHGIE 120  
 YHGHWVGYSD SAVSLSTCSG LRGLIVFENE SVLEPMKSA TNRKYLPFAK KLSKVRGGCG 180  
 LIEIANIHLN KAIVPPFPST WARRHKRETL KATKYQVLI VADNREFQRQ GKDLEKVKQR 240  
 LIEIANIHLN KYRPLNIRIV LVGVVEWDM DKCSVSDPFF TSLHEFLDWR KMKLIPKSH 300  
 DNAQLVSGVY FQGTITGGAP IMSMCTADQS GGIVMDHSDN PLGAAYTLAH LILHINPMHNI 360  
 40 DTLDRCSGCV MKTEKGKGGP LKATGYFFPM VPSRNSKGL ELSLEGGVY CLPLPPIVRE 420  
 SFQSGKQGNR PVIEEGECDC GEPEECNMNC QNATCTILPK DAVCAHGLCC EDQCLKPAGT 480  
 ACRDSSNSCD LPEPCTGASP HCPANVYLHD OHSCQVDVGY CYNGICQTHB QQCVTLWPGF 540  
 AKPAPGICFE RVNSAGDPYG NGCKVSKSSF AKCEMRDAXC GKIQCGGAS RPIVGINAVS 600  
 IETNPLQQG GRILCRGTHV YLGGDDMPDG LVLATGKCAD GKICLNRCQD NISVGVHEBC 660  
 45 AMQCHIGRGV NKNXNCHCSA HWAPPFCDFK OFGSGTDSGF IRQADNQGLI IGLVITLCL 720  
 LAAGFVYVLR RYGLRLLEPT NKHTTIEELR CYRNPPIRG FQYQAHGLR LGKOLMRKPP 780  
 DSYVPKDNFR RLQGCQNDVI SRPLNGLNVP PQSTQRVLP FLHRAPRAPS VFARPLPAKP 840  
 50 ALRQAQGTCK PNPFQKLPA DFLARTITRL HALARTPGW PTEHLRLAPL PAPQYPHQVP 900  
 RSTHTATYIK

Seq ID NO: 11 DNA sequence  
 Nucleic Acid Accession #: NM\_030194  
 Coding sequence: 400-1584 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 GGCAAGAGGG CTGAAGCCCG GAGGAGCGTA GGCAGGGCAG CGCTGGCCCG AGCGKGGACA 60  
 GGAGCGCGCG GACCGGCAAA AATACACGGG AGCGCGTCGC CGAAAAGAGT CCGCGGCTCT 120  
 CTCTGTGTA A CACACTCTCC TCAACCGGCG CTTCCGCCCT CGCTCTGGCG GCGCGCCGCG 180  
 60 TGCGCGCGCG AGCGCGCTCC GACTGCTATG TGAACCGGAG GCTGGGGGAG GAAGGGGACA 240  
 GGGAAAGAAG GGCCTTCCCG CGGAGAGCCT TGAAGACCAA GTTTCGCGCC ACTCTTCGAC 300  
 GCGTCCCTCT TTAGCTCTCG CCGTCCCTCT TCGGACGCT AGCGGGGACA GTTCTCTG 360  
 40 TCTCTCTCG GCGCGCAGCG GCGCTCGGTC CCGGGACCA TGTGAAGAT GAGAGAGCTG 420  
 CCGGAGATGG ACTCGAGTGT GCTCAAAGG CTGATGAACC GGGACGAGAA TGGCGCGCGC 480  
 CGCGCGCGCA GCGCGACCCA CGGCACCTGT GGGCTGCGGA GCGCGCGCAA GTGCGCTGCT 540  
 65 CTGAGCTGCA GAGCTTCTT CGGCGACAG CGCGGGCTACA TCTTAGTTTC GTGCAACGCT 600  
 CGCTGTAAAC CACTGTCCCG GCGCGCGGCT TAGGCTCTCG TGAAGTGGA CAGAGACTCG 660  
 CGCGCGCAGG AGGAGGTACG GCGCGGCTT CGCTCGCGCC TCTACTCGCG GGTCACTGCT 720  
 TACGACGAGC GCAGCGCCGG CGCGGAGAGC CTCGCGAGG ACAGCAAGCT GTGCTGTGTT 780  
 70 TGTCAAGCGC TGGCGCGCAA CGCGGAGGCG ACGGACATCT GCTGTCTCAA AGCGCGCATC 840  
 GAGAGGTTT TCTCGGAGTA CCGAGAATTC TGTTTCTAAA CCAAGGCCCT GGCAGCACTC 900  
 CCACCGCCCG TCCCGGACAG CGGCACAGAG CCGTGGAC CTGGCTGACG CTCTCTCGG 960  
 ACCCGACTAC AGCAGACAGG GGTCTCTG TGAGTACTCT CTCTCTCTCA CTCTCGGAG 1020  
 75 GCTTCACTG CTGCGCGGAG AGACATCTGT GACGCGCTTG GCATCAAGCG TCTGTGTAAT 1080  
 GTCTCTCGG ACTGCGCCAA CCACTTTGAA GGACACTAIC AGTACAAGTG CATCCGACTG 1140  
 GAAGATAACC ACAAGGCGGA CATCAGCTCT TGTTTCAATG AAGCCATAGA GTACATCGAT 1200  
 CGCGTGAAAG ACTCGCGTGG GCGCGTCTGT GTGCACTGG AGCGGGGACT CTGCGGCTG 1260  
 GCGACCATCT GCGTGGCCTA CCGTATGATG AAGAAGCGGG TGAAGCTGGA GAGAGGCTTC 1320

GAGTTCGTTA AGCAGCCCG CAGCATCATC TCGCCAACT TCAGCTTCAT GGGGACGCTG 1380  
 CTCAGCTTGC AGTCCACAGT GCTGGCCACG TCTCTGTGCTG CGAGGGCTGC TAGCCCTCG 1440  
 GGACCCCTGC GGGAGCGGGG CAAGACCCCG GCCACCCCGA CCTCGAGTCT GCTCTCAGC 1500  
 TTTCGGGTCT CCGTGGGCGT GCACTGCGCC CCGCAGAGCC TGCCCTACCT GCACAGCCCC 1560  
 ATCAACAACCT CTCACGCTG TTAGAGCCCG CTRGGGGGCC CLAGAAGCAG AGCTGCTGCC 1620  
 CAGCAAGGCT AGCAGCGCGT GCATGGCGCA GAAAGTTGGG ACTGAGCAGC TGGAGCAGCG 1680  
 CGACCGAGCT CTTCCOCAT CATTTCTCTT TGGCCAAACGA CGAGGCCAGC CAGAATGGCA 1740  
 ATAAGGACTC GGAATACATA ATAAAGGCAA ACAGAAACACT CCAACTYAGA GCAATAACCG 1800  
 GTCCCGCAGC AGCCAGGGAA GACCTTGGTT TGGTTTATGT GTCAAGTTTCA CTTTTCGAT 1860  
 AGAAATTTCT TACCTCATTT TTTAAGCAG TAAGGCTTGA AGTGAATGAAA CCGCAGAGCT 1920  
 CTAGCAAAATG TCGCCACCCA GCTTTACTAA AGGGGGAGGA AGGGAAGGCA AGGGATGAG 1980  
 AAGACAAGTT TCCGAGAATG CCGCTGTCTT GGTGAGTTGT CCGTTGGTGT TGTGTGTGT 2040  
 AGTTAAAGGA ATTTCATTT TAAAAGAAAT CTTGAAAGGT GTGGTTTCA TTTCTCAGTC 2100  
 ACCAACAGAT GAATAATAT GCTTAATAAT AAGATATTA TTAAGCATT CTTCAAGATA 2160  
 TGAAGGTACA AAAAGCTTAG TTACAGTGGG TTTAGAATAT ATTATGTTG ATGTCAAAAC 2220  
 GCTGAGCACC GTAGCATGCA GATGTCAAGG CAGTTAGGAA GTAAATGGTG TCTGTAGAT 2280  
 ATGTGCAAGT TAGCATGATG AGCAACTTGA GTTTGTGCC ACTGAGAAGC AGGGGGGTG 2340  
 CCTGGGAGGA GGAAGAAAGG GAAGAATTAG GTTTGAATTG CTTTAAAAA AAAAAAGAAA 2400  
 AGAAAAAGAC AGCATCTCAC TATGTTGCCA AGGCTCATCT TGAGAAGCAG GCGGGTGGG 2460  
 TGGGAGGAGG AAGAAAGGGA AGAATTAGGT TTGAATTGCT TTTTAAAAA AAAA

Seq ID NO: 12 Protein sequence:  
 Protein Accession #: NP\_001385

1 11 21 31 41 51  
 MVTMEELKEM DCSVLKRLMN RDNFGGAGG SCSHTLGLP SGKCLLLDC RYFLAHSAGY 60  
 ILGISVNYCN TIVSRKGS YSLQLPAB EYVRARLESG LYSAVIVYDE KSPRAESLR 120  
 DSVLSLVVQA LRNAERTDI CLKGGYBFP SSEYEPFCKS TKALAAIFF VPSATPEFLD 180  
 LGCSSCGTPL HDQQGPVEIL PFLYLSGAYH AARRDMLDAL GITALLNVSS DCPNFHFGHY 240  
 QYKCPVEDN HKADISWFM EALBYIDAVK DCRGRVLVHC QAGISRATL CLAYLMKKRK 300  
 VRLEAFEPFV KQRKSISFN PFSMQLQLFQ ESQVLAISCA ARAASFQPL RERKGTPTAP 360  
 TSQVFVFPFV SVGVHSAFSS LPYLHSPIIT SPSC

Seq ID NO: 13 DNA sequence  
 Nucleotide Accession #: none found  
 Coding sequence: 68-340 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 AGCGCCTTGC CTTCTCTTAG GCTTTGAAGC ATTTTGTGCT GTGCTCCCTG ATCTTCAGGT 60  
 CACCACCATG AAGTCTTAG CAGTCTTGT ACTTCTGGGA GTTTCATCT TTTGCTGTC 120  
 TCCCAAGAT CCGACACAGC CTGCTCAGC TGAACAGATT CCAGCTGAGC GTCTGCTGCA 180  
 TGATGAAGCC CCGTAGCTG AAACCACTGC TGTGCAACC ACTGCGACCA CTGCTGCTCC 240  
 TACCACATCA ACCACCGCTG CTTCTACCAC TGCTGTGAAA GACATTCAGC TTTTACCCA 300  
 ATTGGTTGGG GATCTCCGCA ATGTGTAGAT GTGTCCCTGA GATGGAATCA GCTTGAGTCT 360  
 TCTGCAATTG GTTCAACAAT ATTCATGCTT CCGTGTGATT CATCCAACCT ATCACTGCT 420  
 CTACGATATC CCCTTATCT CTAATCAGTT TATTTCTTT CAATAAAAA ATAACATATA 480  
 GCGAGTCAAC AT

Seq ID NO: 14 Protein sequence:  
 Protein Accession #: none found

1 11 21 31 41 51  
 MKFLAVLVLL GVSFLVSAQ NPTTAAPADT YPATUPADDE APDAETTAAA TTATTAAPT 60  
 ATTAASTTAR KDIFVLKFW GDLNPRGVCP

Seq ID NO: 15 DNA sequence  
 Nucleotide Accession #: NM\_016640.2  
 Coding sequence: 38-1358 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 GCTTAAGTGT ACCTCTGGGT CCGGAATGCG GGGGAAAAGT GCGCGCGGCC AGGTGTTGGA 60  
 GCGCTTTGCT ACGCGGTGCG AGGCTTTTAT TGCACACGCG GGTCAATGCC GCGCCACG 120  
 CTACAGAAAC GACCTCCCAA GACGTGGGG CGACCCCCCT CCGCGCGTAC CCGCGGATTG 180  
 TGCGCTCCAT GACAGCGGAC AGCAAAAGCTG CACGGCTGCG GCGGATCGAG CCGTGACAGT 240  
 GCACGGGTGCA CGCTGCGGAG TCGGTGAGCG AGAAGCTGCG AATCTCACC AAGATGCAAT 300  
 TTATGAAGTA CATGTTTAC CCGCAGACCT TGCACATGAA TGCCGACCGC TGGTACCACT 360  
 ACTTCACCAA GACCTGTGTC CTGTGCGGTC TGCAGCGGCC CCCAGCGGAG CCGGAGCCCG 420  
 AGCCCCAACC CCAACGTGAA CCGTGGCTGT ACTTCGCGG GCTGCTGTCG GTGCGTCCG 480  
 ACTGCTGCTC CGAGGAGCAC TTCTACTGCG GCGCAGGCG GCGGCTGCGC CATTACAGAG 540

ABAGCGAGGT CATATCTTTG CCTTCTCTGG ATCAGCTGGT GTCAACCTTC GTGGGCGCTCC 600  
 TCAQCCCA CAACCCGCG CGCGCGCTG CGCGCTGCA TTATGATGCG CCAATTCATT 660  
 TTACTGGGT CGGTGGTGAA GAAATTATTC CTGCTGGTCA TCGAAGAGGT CGAATTGATG 720  
 ACTTGGGATA CCAGATAGAT GATAAACCAA ACAACAGCAT TCGAATATTC AAGCAACTGC 780  
 CAGAGTTTGT GCCATTGGAT TATTCTGTTC CTATAGAAAT CCCACTATA AATGTAAAC 840  
 CAGACAAATC TCCATTATTC AAAOCCGAGT ATGAAAACCA CATATTGTGT GGCTCAAAAA 900  
 CTGCAAGATCC TTGCTGTATG GGTCAACACC AGTTTATCT GTTACTGAC AAAATTAGAA 960  
 GGGAAAGCT TTGGAGGACA AACTGTGCTG ATCAGATAGA AGTTGTTTTT AGAGCTAATG 1020  
 CTATGCGAG CCTTTTGTCT TGGACTGGAG CACAAGCTAT GTATCAAGGA TCTGTGGAGT 1080  
 AAGCAGATGT TACTCGACCT TTGTCTCCCG AGGCTGTGAT CACAGATGGA AAATACTTTT 1140  
 CCTTTTCTGT CTACCAGCTA AATACTTTGG CACTGACTAC ACAAGCTGAT CAAAAATACC 1200  
 CTGCTAAAAA TATATGTGGG GGTACACAAA GTAAGCCTCT TTATGAAACA AATGAGBATA 1260  
 ATGATGTGAA AGGCTGATAT GATGATGTCT TACTGAGAT AGTTCACT GTTCTGATGATA 1320  
 GACCAAAAAA AGAAAAATCA CAGCTGTGGG AAAACTGAAA AAGCATATTT GATTGAGAAC 1380  
 TGTGGGAATA TTAAATTTT ACTGAAGGAA CAATAATGAT GAGATTGTGA ACTGTCAACT 1440  
 ATYAAATACA TTGATTTTTG AGACAAATAT TTCTTATGTC AACCTGTAT TAGACTTCTT 1500  
 ACTGCTGCTCA AATTCTACAC TGAAGAGATT AATTTAGTT ACCTTTTGT GATTTAAAAA 1560  
 TAATTGCATT TGTATATGTC TAAGTGATAA GACAAATTGA GTTATGAGCTT TATTATATGC 1620  
 ACATTTTAAAT ATAAATGCGC AAATGCCAAA TAAATGCTA ACATCTGAAA TTAGTATATT 1680  
 AAAAGAACCC ACTGC

Seq ID NO: 16 Protein sequence:

Protein Accession #: NP\_057724.1

1 11 21 31 41 51  
 I HAAARCRWPL LRGPRLSLRT AANAATATTE TTSQDVAA TP VARYPPIVAS MTADSKAARL 60  
 RRRWRQATV HAAALRPMK NYVPTGT HADRWYQVPT KIVPLDGLP 120  
 PPAEPPEPE PEPPEPALDLA ALRAVACDCL LQHFYLRRL RRVRYEHESE VSLPFLQL 180  
 VSTLVGLSP HNPALAAAL DYRCVPVHFYV VRGHEIPIRG HRRGRIDDLR YQIDDKPNQ 240  
 IRISKQLAEF VPLDYVPIE IPTIKCKPK LPLFKQYEVN HIFVQSKTD PCYGHGTQF 300  
 LLPDLRRER LLRQNCADQI EVVFRANAIA SLFAWTGAQA MYQGFQWSEAD VTRFPVSQAV 360  
 IHDGKYSPF CYQLNTALT TQADNNPKK NICWGTQSKP LYTEDNDV KGFNDVLIQ 420  
 IVHFLNRPK EERKULLEN

Seq ID NO: 17 DNA sequence

Nucleotide Accession #: NM\_025859.1

Coding sequence: 3-2150 (undefined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 GCATTAGCTT GGACTGCACC AGCCATATCG CGCTGGGTGG CGCTTGCCCA GCGCCGGAAG 60  
 AAACCTTAGA CATCTTTTGG GAAGTCCCGG TCACGGGGGA GCAATTAAAC CACTATCGCA 120  
 ATGTGGCTCA AAATGCTCGA AGTGAACCTG CAGCAACTTT GGTCAAATTT GAATGTGCTC 180  
 AGCTGAGCTT TCAAGACCTC CGATCCAAAGA TGCTTTCTAA AGAAGTCTCC TGTCAGAAC 240  
 TGAAGCTGTA AATGGAGAGC TACAAGGAAA ACAATGCCAG AAAATCATCT CTCTTAACCT 300  
 CTTTGAGAGA CAGAGTCTAG GAACATAGAAG AAGAATACAG AGCACTTTCC ACTTTTAAAA 360  
 TCAAGAACCA AATCAACAGA CAGCTGTGCA TCAAGGGAGAA CAGGAPATTA AAGAAGAAAG 420  
 TTGTAGTTT AAATCAAAAA TTCAAAAGGT GTTCAAAAGA AAAATGAGGAA AATGAAGAAC 480  
 AAGTTTCAAA GAATTTGAGG AAACATGAGG AATTCTGAC TCAACTGGCT GACTGTCTGG 540  
 ATCCAGATGA GAGGAATGAC AAGGCATCAG ATGAAGATTT AATTTTAAAG CTTAGAGACC 600  
 TGCCCAAGAA AATGAATCT GTGAAGGAC AAATGTTAT TCTTGAGAG ACTATAAATG 660  
 TCCATGAGAT GGAAGCAAAA GCTAGCAGAG AAACATCAT GAGGCTGTCT TCAAGATGTA 720  
 ACAGAGACGA GAATAAGCT CACTCTGTGA CTGAGAGAAA AGAGAACTGT ACCAGCAAG 780  
 TCTCAGTGTG TTATAGAGCA AAGAAGACTC TTGAAGGGGA AGTTAAGATC TTCCAAGAAA 840  
 GGCTGCTTGC TGGCCAGCAG GTCCTGGGAT CCTCAAGACA GGAAGTGAGC CTCTGAGAAA 900  
 AAGAGCTTCT TGAATGGAG AAGAGTTTGA AGGCCAGTCA GGATGCAGTC ACAACCTCAC 960  
 AAAGCCAGTA CTCTCATTT AGGGAGAAAA TCGCAGCCTT ACCTTAGGGC AGATTGAGCA 1020  
 TGAAGTGGCT CACTGAGAGC ACCTATTTTG AGAAGATCTG AGAATGAGC AGCGGGAG 1080  
 AAAGCAGGGA CCGGATGGTC TCGAGCTTGA AGGCCAAT ATCTGAGCTT GTTGAACAGT 1140  
 TGGGAAGAGA GTCTGGGTTT CACCAGAAAG CTTCTCAGAG GGCCCAAGAAA CGAGAGATA 1200  
 TGTGGAAGAC TCTTCAAGGT CAGCTGACAC ACTGTGAGGC AGAGCTGGTT TCTGGAAGTT 1260  
 TTTTGGGAGA CAACCTGAT TTTGAGAAAC AAAAATATCT TAAATTTCTG GATCAGCTTT 1320  
 CTCAGAAAT AGAGTTGGAC CAGATGGCTG CCGAAGCTGG CTTTGACAC CGGCTGGAG 1380  
 TGGTTTAGC TGAAGAGAG CAGCTGGTTC GTTGTGAGAG CAATGCACT TTTGAGAAC 1440  
 AGACCAATTC CCACAATTG CAGAGAAAGC TAAAGACACA GAAGAGAGAGA CTGGAGAGCA 1500  
 AAGAATTACA CATGAGCCTC CTCGCGCAGA AAATAGGCCA GCTGGAGGAG GAGAAGCAGG 1560  
 CACCCACAGC CTTGGTGGT GAGAGGGACA ACBGCATCTT TACCATCAGG AACCTGCAGA 1620  
 AGAAGGTGGA GAGGTGGTCA AAGAGCTGGA ACAAGTGTGC AGACTGTGAC ACCGAGGTCA 1680  
 AAGCCAAATC GCGGAGACA AATGAACTGA AGATTAACCT TTGGACAGAG ACTAAAGCCA 1740  
 TTGAAGATCT AAACCAATCC AGAGACCAAC TGGGAAGAT GAAGGAGAAA GCTGAGAAAA 1800  
 AGCTCATGTC TTCTAAAGTA GAAGTGATA CCACAGAAACA TGAGGCTAAG GAGAAATAAG 1860  
 AAAGGGCCAG AAACATGATA GAAATGGTAA CCAAGTAAAT GAAGACACTA AAAAATCTCT 1920  
 TGGAAAGAGC AGAAAGAGA GAAAGCAUC TGGCAGACTT CAGGGAGGTT GTGTGCGACA 1980  
 TGCTAGGCTT GAACGTGACC AGCTTGTCT TTCTGATTA TGAATGATC AAGTGTCTTG 2040

AAAGATTGGT CCATTCACAT CAGCATCACT TTGTACCTG TGCTGCGCTC AAAATGTGGA 2100  
 CTACTGGGCA AGAGAGGGCAC CCACAAGGCC ATTACAGCT CTCTCATGSA ACACCTGATC 2160  
 TCTTGAAGGA GTGTGGCATA AGACATUGGA CACAATTCCC AATTCCAGAA ATTCCTCATG 2220  
 TCTTGGAGAT TTGATCAGT TTGGAATATT TTATGCTTTG ATGATATAGT GAGATGCAAT 2280  
 CACTTGCAAA AACGATCTCA AAGGTGTGAG CCTTAGATAA ACGTCAGCAT TAAAAAAGCG 2340  
 CAAAAAAGAA AAAAAAAGC ATTTAGGAT CCAGAGAAGT TCCACAGAT TGCATGAGTT 2400  
 AGATTGGGAA ATGGAGTTGG GAGATAAAT TGGGAGGAT CTATTATTAG TCAGGGGCTT 2460  
 TACTAGCGGA TTGATGCTG ACAATAAGCA TTGTGGAGAG CTGTGACATT TTAAATTGAT 2520  
 AACCTTCTG GGGCTGAGAC ATAAAGTTAC CTATCCAAGG TTGAGTTGG GTAGTTGGTG 2580  
 GACCAGGATG GACAACACTT TGGCCGTGCC TCAAAAAGCA TACCCTCTCT CCGTCTATCG 2640  
 AGAATCTGTT TCCATCGAAT CTCTGTGATG CTGGTGGGAA TTGTTTGCAT AGAGGAAGGA 2700  
 CAATAACCCCT GCGATCGTGA GTTAATGTCC GGGCTGTGCA CAGTGTGTCA TGGCTGTAAT 2760  
 CCGACGACT TTGGAGTGCA AGCAGCGCAT ATCATTTTAT GTACAGGATT TAAGACCAAG 2820  
 CTGGCTAACA TAGTAGAGCC CTGTTTCTAC TAAAAATACA AAAATAAGCC AGGTGTGGTG 2880  
 GTGCATGACT GTAATCCCAA CTACTCAGCA GGAGAAGCAC TTGACCCAG GAGACGGAGG 2940  
 CCGCAGTGAG CCAAGATTGT GCCACTGCA CCGAGCGCTGG GCGACAGAGT GAAACTACAT 3000  
 CAGAGAAAAA AAAAAAAGAA AAAAAAAGAA AAGTTAATGT CCAAAATGCA CAGATTACCA 3060  
 AGTGAAGCT ATATGATTCT TTCAAAAAGC AAAAGCAATA TACCTAATTC ATTGGATCA 3120  
 AACCTACATA GGTCTAGGT CCGTAAGAA ACTTGGCTGT TCTAACGTTT GCTACAGAT 3180  
 TATATCTGGT GGTAAATGTT AATGTTTAC CAGGCGTGT CTCACCTCT TAAATGGA 3240  
 AGCTTTATT GTAGGCCAAC CCTATAAGAT GAAGAAATCC TATATAGTCT TATTCACCAA 3300  
 TATATCCAAT ACACCCACAG CAATGGTACC TTTTAAGAT CAGGATTTTA TTATTAATTC 3360  
 CTGTGACTTT CTGTTTCCA TTAAATTTC TATTTAGAA ATTTTTCAG GAATCATATT 3420  
 AACTTACATA GTCTCAGGT CCGTAAGAA ACTTGGCTGT TCTAACGTTT TCTATCATCA 3480  
 TAAGATGAT TTTTTATTG TCGTTAAAG AAGCTCTAGC ATGAAATPAA AAGGAAGGGA 3540  
 AAGAAATAGA TCTGTGTCAC CGGAACATTA GGAGAAATG AAAAAATAC AACCAACCGT 3600  
 TCGTGATGCA TCAAAAAGTC AAGTCAGCC TGGCCAACAT GGCAAAACCT CCGTCTGCA 3660  
 AAAAAATACA AAGTAGGCC GGTATGTTGG CATATGCTTC TAATCCGAG TACTGAGAG 3720  
 CTTGAGACAA AAAAAATGCT TGAAGAGCG AGGGCAGAT TCAATGAGC GAGAGAGGAG 3780  
 CTACTGCACT CGACCTGGG CAACAGAGAG AGACCTGTC TCAAAAACCA AACCAACCAA 3840  
 AAGGTCAAG TCATAATAAG CAAATATTG CGTCTTTCT TCTAGACVAA AAGAAATPAA 3900  
 AGAGATGAAA CAATCAATTG CAAGGGTCAA AACTAGATTG GATCTGGTT TGAATGAAA 3960  
 AAGACATAAA ATATCTCTGC AATAATGTA AAAAAATGAA TTGGGACTAA GTCTCATGAT 4020  
 ATATTAAAT ATTTTAAAT TTTAAGCTT GACAAATGCA CTGATGTTA TACITTAAT 4080  
 AACTAAAT CTGAAATCC ACACTGCTAC AGACAATAA TGATAAAATG GGAATAAATA 4140  
 AAAAAAATA A

Seq ID NO: 18 Protein sequence:  
 Protein Accession #: NP\_079335.1

1 11 21 31 41 51  
 MSLDCTSHIA LGAASPAPEE TYDHLSEVPV TREQINHYRN VAQNARSELA ATLKVPEACQ 60  
 SBLQLRSKM LKEVSCQEL KAMESYKEN NAKSSLLTS LDRVRQBLEE EBAALSTK 120  
 RRTTAAAT KENQELKKV VLNELKQKG SKINENKKQ YSKNCRKHEE FLTQLRDLCD 180  
 PDERNDKASD EDLILXRDL RKENEVFKQG FVLEETINV HEMEAASRE TMRLESEVN 240  
 REQKKAASCT BEKELNLQDL LSAVEAKEAL EREVKPIQR LLAGQQVWDA SKQEVSLKK 300  
 SSSHLEKSLK ASQDAVTSG SQYSEFREKI AALLRGLRM TGSTEDTLE KIREMDSRBS 360  
 SKDRMVSQLE AQISELVEQL QKISGHQGA LQKAGAKENM LETLOGLLTH LEALVSGV 420  
 LRDNASEQSL KRLDNLQSL QKMLDQKAA BLQGTGGLVT VLAKETLQVY LSNVAENK 480  
 TIAHNRQKL TKQERLESE ELHMSLLRQK LQLEBEKQA RTALVVERDN AHLTNLQK 540  
 KVERLQKELN TCRDLHTELK AKLADTNELK IKTLPTQKAI EDLNLNRDLQ EKMEKEAEK 600  
 LMSVKSLBDT TEHSAENKE IARNMIEVVT SEMKTLKSL EBAKSRERQL ADPREVVSM 660  
 LGLNVTSLAL PDYENKCLE RLNVHSIHQHF VTCACTKIDVT TGQERHPQHL LQLLH

Seq ID NO: 19 DNA sequence  
 Nucleic Acid Accession #: AF071552, NM\_000602  
 Coding sequence: 441-1313 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CTTTGTATAA GGCTCAGCTA AAAGGGAAAT TGAGTGGGTC AGGTACACCG GATACTATAC 60  
 ACTCTATTGC ATGATCTCTC TGCTCATGAC AGAAGACGTT TATAAGCTTA TTTAAAGGA 120  
 TACCAAGTTG AATGATCTCT TTATTATACA CCAAGAGAAG CATGAAACAG CTGTTATCA 180  
 TTGACTCAT CATTTAATCT GTATTCCAG CTCTTCACAC TTGAAAGAAG ACATAATCA 240  
 TTCTTCACAG GATTCTGGGA CTATTAACTG AACTTATGTG TGTAAAAAGGA ATTCATACA 300  
 TGAAGACAGT AGAAATAAT ATTACTACTA TAACCAATTGT ATTTTACAT GTTTAAATA 360  
 TAGCCATAAT TAGCTACTAC AATTCCAAGT GTAAAAGTAA AATGATTTGC TTTCGTTTG 420  
 TTTCCTTTCG TTAGGGGATC AGGCAATG AGCATATCT TGAAGAATTT GGCTATAGA 480  
 AGCTAGAGAA CTTAAATGCG TTGGAACAT TAACGTGACAT TCTTCAACAC CAGATCCAG 540  
 CTGTCTCCTT TGAAAGACCT AACATCATG GTGGGAGTGC CATGGACTTA GGCTTAGAGG 600  
 CCATTTTTGA TCAAGTTTGT AGAAGAAATC GGGGTGAGAT GTGTCTCCAG GTCAATCATC 660  
 TTCTGTACTG GGGCTGACC ACTATTGGTT TTGAGACCAC GATGTGGGA GGGTATGTT 720  
 ACAGCATCTC AGCCAAAAA TACAGCACTG GCATGATTC CTTCTCTGT CAGGTGACCA 780

TTGATGGCAG GAACACATT CTCGATCTG GGTITGGACD CTCATACCAG ATGTGGCAGC 840  
 CTCGTGGAGT AATTCTGGG AAGGATCAGC CTCAGGTGGC TTGTGTCCTC GTTITGACGG 900  
 AAGAGAATGG ATTCTG0AT CTAGAACAAA TCAGAAGGGA ACAGATACAT CCAAAATGAAG 960  
 AATTCTTCA TCTGTATCTC CTAGAAGACA GCAAAATACG AAAAACTAC TCTTITACTC 1020  
 TTAAGCCTCG AACAATTGAA GATTTTGAAT CTATGAAATC ATACTGACAG ACATCTCCAT 1080  
 CATCTGTGTT TACTGTAAA TCATTTGTGT CCTTCCAGAC CCAAGATGGG GTTCACTGTT 1140  
 TGGTGGGCTT CACCCTCAC CATAGGAGAT TCAATTATAA GGACAATACA GATCTAATAG 1200  
 AGTTACAGAC TCTGATGAG GAAGAAATAG AAAAAAGTCT GAAAAATATA TTTAATATT 1260  
 CCTTCCAGAG AAAGCTTGTG CCAAAACATG GTGATAGAT TTTTACTATT TAAAGTAAGG 1320  
 AGTAAACAAA TCTTGTCTAT TTGTCAATCCA GCTCACCAGT TATCAACTGA CCAACTATCA 1380  
 TGATATCTCT GTACATTTGA CTATTTTGA AAAAAATCTT AGACATAAA TCATTTACCC 1440  
 TATAAAAAAT TCATCATATA TAATTAACAA GCITTTTAAA GAAACATAAC CACAACCTT 1500  
 TCAATAATAT AATAATAATA ATAATAATAA ATGTATTTTA AGATGGGCT GTGTGTATCT 1560  
 TGGAAATTGG TGATTTATGC TAGAAAGCTT TTAATGTGGG TTATTTGTTG AATTC

Seq ID NO: 20 Protein sequence:  
 Protein Accession #: NP\_006653.1

1 11 21 31 41 51  
 MIDRAYLERI GYKSKRNKLD LETLTDLQH QIRAVPPHNL NIHCDDAMD LGLEAPDQVV 60  
 RNRNGWGCLQ VNHLIYWALT TIGFETMLG GYVYSTPAKK YSTGMIHLL QVTDGRNYI 120  
 VDAGGFRSQQ MWQPLELISG KDQFQVPCVF RLTEBNGFVY LDIQRRESQYI PNBEFLHSLD 180  
 LBSYKRYKIY SFTLKRPTIE DFESEMTNLYQ TSFSSVFSK SPCLSLQPDG VHCLVGPTLT 240  
 HRRFNYKDTN DLIEFKTLE SEERIEKVLKNI FNISLQKLV FKHGDRFFTI

Seq ID NO: 21 DNA sequence  
 Nucleic Acid Accession #: NM\_003714  
 Coding sequence: 123-1031 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CGGCACGAGC AAAAAGGAAG AGTGGGAGGA GGAGGGGAAG CGCGCAAGGA GGAAGAGGAG 60  
 GAGGAGGAAG AGGGGAAGCAG AAAGGATGCC GTCTCCCGCA CGGGAGGTTA ATACCAAGAA 120  
 CCAATGTGTC CGACGGCGTC GGCACGTTCA TGACCGTGCC TTGTGTGTGT GCACACTTGG 180  
 ACCCGCGCCG GGGGACCGAC GCGACCAACG CACCGGAGGG TTCGCAAGAC AGAGCACTCC 240  
 AGCAGAAAGG CCGCTGTCTC CTCGAGAATA CAGCGAGAT CTACGACTGT TTGTGTAACG 300  
 CTGGCGATGT GGGGTGTGCC GTGTTTGAAT GTTTCAGAAA CAATCTGTGT GAGATTCCGG 360  
 GCATTACATG GATTTCATG ACTTTTCTGC ACAACGCTGG AAAATTIGAT GCCACGGGCA 420  
 AGTCATTAC CAAAGACGCC TTGAAATGTA AGGCCAGCCG TCTGCGGCAC AGTTCGCGCT 480  
 GCATAAGCCG GAATGCCCG GCATCATCGG AATGTGTGTC CAGTTTCGAC CGGGAATCTC 540  
 AOTCTACAGA CGACCTGTGC GGGGTGTGCC AGGAGAACAC CCGGTGTGTA GTGAGAGTGA 600  
 TCAATTTCAA GACACTGTCT CTGCAAGAAC CTACTGTGGA CCTGTGTGAAC TTGCTGTGCA 660  
 CCTGTGGGGA GGAGGTGAAG GAGGCGATCA CCCACAGCGT CGAGGTTACG TTGTAGCGAG 720  
 ACTGGGGAAG CCTGTGCTCC ATCTGTGAGCT TCTGCACTCG GGCCATCCAG AAGCCTCCCA 780  
 CGCGGCCGCC CGACGCCACG CCGCAGGTGG ACAGAACCA GCTCTCCAGG GCCCACAGG 840  
 GGGAGACGAG ACATCTGCTC CCAGAGCCCA CTAAGAGGGA GACTGGGCCA GGTGCGCAAGG 900  
 GTGAGCBAAG TAGCAAGAGC CACCCAAACG CCGATGCCCG AGGCAGAGTC GGGGGCCCTG 960  
 GGCGTCCAGG ACCITCCGGA AGCAGCGAGT GGGGAAGACGA ACAGTCTGAG TATCTGTATA 1020  
 TCCGAGAGTG AATAGAAAGG CCTGGCCACG AATCTTCTC TCCACGCCGT CCAATTCTT 1080  
 ATCTATGAC ATTCACAAAC ATTACACATT AGAGAGCGGG GATGTGCAC CGAGGATTTCT 1140  
 GTGGGACTGT TGGACTTCT CAGGTGTGTT GTTCGCGGGA CGACAGAGTG AGATGGAGAC 1200  
 CCTTGGGCCG GTGGGGTCTC AGGGGTGCGT GGTGAATTC TCACTTACAC GTACTCAAG 1260  
 GAGGCGGCCG GGGTATCTCT GGTACCTTTG TCTCTTTC ATCTGTGGAG TCAGTGGGTT 1320  
 TCGGCGCGCT TGTGTGGG GAGGTGAACC AGGGAGGGGC AGGGCAAGCG AGGGGCCCA 1380  
 GAGCTGGGCG ACACAGTGGG TCGTGGGCGT GCGGCCGAAG CTCTGTGGTC AGCAGGCTCT 1440  
 GGTCTGTCT CCGCAAGAGT CAGGGCGGCT GAGTCCAGG ACAGAGAGTA ATGTAAAAAT 1500  
 AATATTCCT TAGAATCCAC GAGAAGGCTG GAGAGAGGCG AGGGGCCGAG GGGGTCTG 1560  
 GTGCACAACT GAAATTCAGT TCTTGTGTGT GGGCGTTCGG GTTCAGAGCT CTGGCGGAGG 1620  
 GTGGAGGGAG GAGTGTCAAT TCTATGTGTA ATTTCTGAGC CATTTACTG TCTGGGCTCG 1680  
 GGGGACACAT GTCCAAAGGA GTGGGCCCTA TGAATTTATA TTTTAAACAC TGCTTCAAAAT 1740  
 CTGATTTTCA CTTTITAT TTATCCAGTT ATAICTACAT ATCTGTATC TAAATAAAT 1800  
 GCTTTCAAC AAAAAAAAAA AAAAAAAAAA AAAAAA

Seq ID NO: 22 Protein sequence:  
 Protein Accession #: NP\_003705

1 11 21 31 41 51  
 MCAERLQGM TIALVLATPD PARGTDATNP PEQPDQRSS QKGRLSLQNT AEHQCLYNA 60  
 IDVCGGVPEC FENNSCEIRG LHGICMTPLH NAGKPDAGQK SFKDALCKK AHALRIHFGC 120  
 TNRKCPAIRE MYSVLQYKIH KKHDLCAAAQ ENYRIVIVEM HFKDILLHP YVYLVNLLT 180  
 CGEEYKATIT ISKQVQCHQN WGSLSLSLT CSTAKGKPT APFRQVQD RYKSRALHIG 240  
 EAGEHLPEFS SHETGNGAKG ERGSKSHPNH HARGRVGGLG AQQFSGSSEW IDEQSEYSDI 300



RR

Seq ID NO: 23 DNA sequence

Nucleic Acid Accession #: NM\_05264.1

Coding sequence: 537-1954 (underlined sequences correspond to start and stop codons)

```

1      11      21      31      41      51
|      |      |      |      |      |
5  GAATTCGGCG CAGAGAAAT CTGGGCTGGG AACAGGCCAT TCTCGCGCGC GCCTCCAATA 60
   ACCATTAACA TCCTCAACGA GCATCGGAGC CGAGGGGCTCT GCTCGGAAAT CGTCCTGGCC 120
   CAACCTGGCC CTTCGAGCTC TCGAAGATTA CCGCATCTAT TTTTITTC TTTTTTCT 180
   TTCTCTAGCG CAGATAAACT GAGCGCGGAA AGGGAAGGAG GGGCGGGGA CACCATTTGCC 240
   CTGAAAGAAT AAATAAGTAA ATAAACAAC ACCTGGCTCTCG CGGCACGTGG AGCGGTGGG 300
   CTGAAGTCAG GTTGGGTGCG ACCTGAACCC CTAAAGCGG AACCGCTCC CGCGCTGCC 360
   ATCCGGGAGC TGAATGCGCG GCGGCGGTGG CTGCTGCGAG ACCGCGAGGT TCGCTTTTGA 420
   CTGGATGGAG CTGAACCTTG GCGCGCCAGA GCAGCACAGC TGTCCGGGA TCGCTGCAAG 480
   CTGAGCTCCC TCGGCAAGAC CAGCGCGCGG CTGCGGATTT TTTGGGGGG GCGGGGACCA 540
   GCGCCGCGCC GGCACCATCT TCTGCGGAC CCGTACTTCT GCGCTGCCGC TCTTGGACT 600
   GCTCTGTGCG GCGCAAGTGA GCGGCGGAGA CCGCTGTGAT TCGGTGAAG CCGATGTATCA 660
   GTCCCTGAAG GAGCAGAGCT GCAGCACCAA GTACCCACCA CTAAAGCAAT GCGTGGCG 720
   CAGGAGAAC AACCTCAGCC TGGCATCCGG CCGTAGGCGC AAGGATGAT GCGCGAGCGC 780
   CATGGAGGCC CTGAAGCAGA AGTCGCTCTA CAACTGCCCG TGCAAGCGGG GTATGAAGAA 840
   GGAGAAGAAC TGGCTGGCA TTACTGGAG CATGTACCAG AGCCTGCAAG GAAATGATCT 900
   GCTTGGAGAT TCCCATATG AACCATGTAA CAGCAGATTG TCAAGATAT TCCGGGTGGT 960
   GCGATTCATA TCAGATGTTT TCGACAAGT GAGAGCATT CTCAAAAGGA ACACTGGCT 1020
   GGTACGAGGG AAGCGTCCA ACCTGACCA CATTTGCAAG AAGTACAGGT GCGGTACAT 1080
   CACCCCGTGC ACCACCGAGG TGTCCAAGGA TGCTGCAAC GCGCGCAAGT GCCACAAGCG 1140
   CTCCTCGGAG TCTTTTACA AGGTCCCGGC CAAGCACAGC TAGCGAATGC TCTCTGCTC 1200
   CTGCGCGGAG ATCCGCTGCA CAGAGCGGAG GCGACAGACC ATCGTCCGTG TGTGCTCTTA 1260
   TGAAGAGAGG GAGAAGCCCA ACTGTTTGA TTGACAGGAC TCGTCAAGA GAATTATCAT 1320
   CTGCGATCTC CGATCTGCGA ATTTTCTGAC CACTGCGAG CAGAGATCAA GCTCTGTGAC 1380
   CAGCTGTCTA AAGGAAAACT ACGCTGACTC CCTCTGCGCC TACTCGGGGC TTATTGGCAC 1440
   AGTCATGACC CCGCACTACA TAGACTCCAG TAGCGTCAGT GTGGCCCAT GGTGTGACT 1500
   CAGCAACAGT GGGAAAGACC TAGAAGAGTG CTGGAATTT TGAATTTCT TCAAGGACAA 1560
   TCAGTGTCTT AAAAATGCAA TCAAGCTT TGGCAATGCG TCGATGTGSA CCGTGTGSKA 1620
   GGCAGGCTTC CTGCGCTTTC CTAGCAGAG CACTACAGC ACTGCGCTCC GGGTTAAGAA 1680
   CAAGCTCCG GGGCAGCAG GGTCTGAGAA TGAATTTCC ACTCATGTT TGCCACGCTG 1740
   TGCAGATTTA CAGGCACAGA AGCTGAAATC CAATGTGTG GGCATACAC ACCTCTGTAT 1800
   TTCCAAATGT AATTATGAAA AAGAAGGTC TGGTGCTTCC AGCCACATAA CCACAAAATC 1860
   AATGGCTGCT CCTCCAAGCT GTGGTCTGAG CCGACTGCTG GTCTGTGTGG TAACCGCTCT 1920
   GTCCAGGCTTA TTATCTTTAA CAGAAACATC ATAGCTGCAT TAAAAAAATA CAATATGAC 1980
   ATGTAAAG ACACAAACCA AGTTATCTGT TCTGTCTCT CTGTATAGCG TGAATTTCCA 2040
   GTTAGGAGC TCAGTTGAGA AACAGTTCCA TCAACTGGA ACATTTT TTTTCTTCT 2100
   AAGAAAGCTT CTGTGATCG TTGCGGGCTT CTGTGAAAAA CTGTATGAG TGTCTCATCT 2160
   AAACCTAGAA GCGTTTGGGA TATGCTGAT TTAAAGGGA CAGTTTGTAA CTGCGCTGT 2220
   AAAGCAAACT GGGCGTGT TTGATGATGAT GATGATCATC ATGATCATGA TGATTTTAAC 2280
   AGTTTATCT CTGCGCTTTC CTAGCAGAG AAGAGGATG TTTCTTCTAG GTATGCCAA 2340
   TATCTGCTT ATGACATGT ATTCTAATG ATATAAATT CAGCTACAT TGAIGGCCAA 2400
   CTTTTGTGCG ACAAAAGA TAATTACAA GAGTGGGCTT TGTGGAACA GCTGTGACTG 2460
   ATGTTCACTT TTATATATG ACTAGCATTT TCCAGCTGTA TGTTTATGA CTGTAACAG 2520
   TCTGCACTC TTGTACAAAA GAAAAACCA CCGGGAATTC

```

Seq ID NO: 24 Protein sequence:

Protein Accession #: NP\_005255

```

1      11      21      31      41      51
|      |      |      |      |      |
60 MFLATLYFAL PLLDLLSAR VSGGDRLDVC KASDQCLKBQ SCSTKYRILR QCVGAKETNF 60
   SLASGLEAKD ECRSAMEALK QKSLYNCRCK RGMKMKKNCL RYVWSMYQSL QGNLDLEDSP 120
   SVSNVNRSLD IFRVVPFISD VFQGVBIHPK GNNDLDAKA CNLDIDCKKY RSAVITPCTT 180
   SVSNVDCNRR KCHKALRQFF DKVPKAKHSYG MLFPCSRDIA CTERARQTV PVSVEBEKE 240
   PNCLNLQDSC KTYNYCRSLR ADFPTNCQPE SRVSVSLCKE NYADCLAYS GLGIVMTPTN 300
65 VYDSSLSVA PWCDSNSGN DLGELKFLM FPKENCTLEN ARQARGNSD VTNVQARPV 360
   QTITATTTTA LRVKNCKPLG AGSENIHPTL VLPFCANLQA QKLKSNVSGN TRLCSNGNY 420
   EKEGLGASSH ITKSMMAFP SCGLSPLVL VVTALSTLS LTETS

```

Seq ID NO: 25 DNA sequence

Nucleic Acid Accession #: FGENESH predicted

Coding sequence: 1576 (entire sequence represents open reading frame)

```

1      11      21      31      41      51
|      |      |      |      |      |
75 CTTTGTTC GCCATGGCTA GTCTAGTGGT ATCTGGAATA ATGGAAGAAA ATGGGGGCTT 60
   TGAAGAAGTA GCATGTTTGG GGGGAAGGCG TAAGBACCGA GGGCTGCTGG AGACGAGCG 120
   CGCCCTTCAG CTGGCTCTCG ATCAACTCTG CCTCTGGGT TTGGGGGAGC CCCCAGCCCC 180

```

CAGGCGGGG GAGGACGGG GAGGTGGGG GGGGGGGGG CCCGGGAGC GACAGCCCC 240  
 CCCCAGGG GGGGGGGG GGGGGGGG GGGGGGGG GGGGGGGG GGGGGGGG 300  
 CCGAGCGCAG AGCGCCGAG CCCCACGCG CCCCAGGG GCGAGGAGC CCAAGCTCT 360  
 CGCTCTCTAC AAAGAGGGCG AGCTGGCGCT GAAGGGGAGC AGCAACACCA CGGAGTGTGT 420  
 TCCGTGGCCC ACCTCGGAGC ACCTGGCCGA GATCTGGGC AGGCAAGGCT GCAAGATTAA 480  
 GGGCTGTAGG GCGAAGACCA ACACCTACAT CAAGACACCG GTGAGGGGG AGGAACCACT 540  
 GTTCATGTGG ACGAGGGCAG GCGAGGAGCT GCGCACAGCG CGCGGGGAA TCACTCGAGC 600  
 ACCGAGACAC TTCTCCATGA TCGGTGGCT CCAGCAACAG TCAAGGGGGC CTCTTGTGT 660  
 GGTCTCTGCT CTGCGGGGCG AGGTGACCAT CCGTGTGGGG GTGCGCTACC GCGTGTGGG 720  
 GCTGTGTGGT GGGCCCAAG GGGCAACCAT CAAGCGGCAT CAGCAGCAAA CCAACACATA 780  
 CATTATCACA CCAAGCGGTG ACCCGAGGCC CGTGTTCGAG ATCACGGGTG CCGCAGGCAA 840  
 CGTGGAGCGT GCGGCGAGAG AGATCGAGAG GCACATGGGG GTGCGCACTG GCGAGATCT 900  
 CGAGTACAA AGTGAACAAC ACTCTCGGG GGGGAGCGCG GACGAGCAA TGGATAGCGC 960  
 CTACTCCGAC GCGTGGGGGG TGACACGAGC CGGTGCAAG CCGCTCTCCA CTCTTCGGCA 1020  
 GAACAGCCTG GCGTGCATG GCGAGTGGG AGTGGACTCT GCGTTTGAAG CCGCAGCGCT 1080  
 GGGTGGAGC GGGCGGGACT TTGGCTACGG CGGGTACCTC TTTCGGGGCT ATGGCGTGGG 1140  
 CAAGCAGAGT GTTATCTACI GCGTGGCCGA GACTAGGCCG CGCTGTGGG CGGCGCAGGA 1200  
 GAAGGGCAGC GCGAGTCTCT TGCTCTCTC CTCTGGCTCC TCCCTCTCT CTCTTGGG 1260  
 CAAGGCGGCG GCTGGGCGCG CGGGGACACA CCGCTCCCTT GCGACTTCC GGGGACCGGA 1320  
 GCTGGCGGGA CTCGCGAGGC GCGCGCGGG AGAGCGCGCT CAGGGCTTCT CTAAACTTGG 1380  
 GTGGGGCGGC CTGCGGAGGC CGCGCGGGCG GCGGATATGC ATGTGCTGCT TTGAGAGCGA 1440  
 AGTGACTGCC GCGCTTGTGC CCGTGGGACA CAACCTGTTC TGCACTGGAGT GTGCACTACG 1500  
 CACTCTCGAG AGGACGGAGC CAGAGTGTCT GGTCTGCCAC ATCACAGCCA GCGAAGCCAT 1560  
 CCGAATTTC TCTTAA

Seq ID NO: 26 Protein sequence  
 Protein Accession #: P01BNH predicted

1 11 21 31 41 51  
 PCFAMPSLVV SVIMERNGGF GHI GCPGSGA KDRGLLEDER ALQLALDQLC LGLGEPFAP 60  
 RAGEDGGGG GGAPAQPTAP PQPAPPPPA APFAAPTAP AAQTGPQPTA PKGASAKLC 120  
 ALVFAELRL GCGAGTTCV PFTSEHVAV IVBGQCKCK ALRAKINTYI KTVNGEPPY 180  
 FMVYGRBDEY ATARREISA AEHFSMIRAS RNKSGAOPV APALPGQVYI RVRVPRVYG 240  
 LVVPGKGTAT KRIQQQNTNYI HTPSRDNDP VFETTGAPGN VYARREIBIT HIAVRTGKL 300  
 EYNNENDFLA GSPDAIDSR YSDAWRVHQH GCKPLSTRQ NSLHGICGCG VDSGEAPRL 360  
 GHQGGDFGYG GYLFPQYGYG QQVYVYGVAE TSPPLWAGGE NATFISVLFS SASSSSSSA 420  
 KARAQPGAH RPSATSGAPE LAGLIRPPFG BPLQFSKLG GGLRFSFGG KDCMVFPISE 480  
 VTAALVPCGH NLPCMECAVR ICERTDFPC VCHITATQAI RPS

Seq ID NO: 27 DNA sequence  
 Nucleic Acid Accession #: P01BNH predicted  
 Coding sequence: 1-2070 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 ATGAGCGGTG CCGGGGGTGC GGTGGGAGC CCGCCCCCA GCTCGCGGAC CCGGGGCTCT 60  
 CGGGCGCGCG GCGAGCGCCC CTCTGTGGCG GTGCAATCTT TGAAGCGCGA GACGCCGAC 120  
 CTCAGCGAGA GCGAGTTCGA GAAAGGAGC CTGCGACTGG AGAAAGGCTC GAGTCTCTG 180  
 CAGCAGCAGC ACTCGGAGAT GCTGGGCAAG CTCCTATGAG AGATCAGACA TCTTAAAGCGG 240  
 GAAACAACAG GTGAGCGGCG GCGGGGCGCT AGCGCGCGCC TGCGTCCCA GCGCAACTCA 300  
 ACACTGCGCG TCGCGCAGCA CAGAACAACA GCGTCAACT CCAGCAGACG CTGGGCTGCA 360  
 GGGGGAACAC AGGAAGGGGA GCGCTCCAG ACTGTGCTTG CCGACTGGG TGCACTGGCC 420  
 CGTGTATGCC AACCAGATGG GTACAGGTTG TGGGGAGCT GCGACGAGCT GCGTACTCCT 480  
 ACGCGTGGCT GCGAGTATT ATGACGCCAA GACAGCAGCG TCGAGTCTCT GGGAGAGGCCA 540  
 GGGCTGAGG TCAATTGACG GCGGCAAGTG GCGACAGGHT GCTCCCGAGA CCGCTCTCT 600  
 CCAAGTAGAG CTGAATAGG AAGGAACCC TGGGACAGCC CCGTGGCTTG TAGATCTTGG 660  
 CTCGAGATTG CTGCTTGGG CAGGCCAGG ATTTCAGGC CTATGGCTCT GAGTCTTCAC 720  
 ATGCTGGGGG CCGAGGGGAT ATGGACACAT CTGATCAGG GATCCCTTCC TGCACTGGG 780  
 GCACGACA TGGGGACAAA GGGAGGAGAG AGAGTCTGCT TTCTCTGCA CTGTTCGAG 840  
 GCGACTTCCC ATCTGTACAG GGGGCCACG CCGAGCCAGG ATCTGGGCT GTGGTCTCAA 900  
 GCTCACTTCC CATTATCTTT GGGGCTGGG CTGACATCAG GAGGACATCT GACTGTGTGA 960  
 TGTAGCCAGC CTGGGAACAT GCGAGCTGG GCGATGCTTA GGGCTCTCC TTCCGAGGGA 1020  
 GACATGGA GAAGGGGTGA GGGAGGGGCC TTGCTATGCC CTGTGAGCA CTCGATGAG 1080  
 CTGTCTTGGG CAAAGTGTGG CCAAGTGGG GAGGCCAGC CTGCAAGTGC TGGGGAGCT 1140  
 GACGAGAC GGAAGAGGC CAGTCTTCT CTGGGAGCT GCTGTTCCT GTGTCCCAAG 1200  
 CCGCTCTGCT TTGAGATGG CCGCTCAGGA AACCACCTTT CCAGGGGCTC TGCTCCCTTG 1260  
 GGGCTGCT GGTCTGCAT CAACGGAGTG TGGGTAGAG CGGGAGGACC CAGCGCTCCG 1320  
 AGGCTGAAG AGGCTGCTCT ACAGACACAC AGGCGAGGAG GCAAGGGTGG GCGTCTTGGG 1380  
 GCGGTAGCG CCGAACACTGT GGGTCTCTCT GCGACAGCTC TCTCAATGTC AAGCTCGAG 1440  
 TCTGTAAAT CCGTCACTCT TTGAGCCAA TCTCAAGCA AGGCGAGCG CCGAGCGCGG 1500  
 TCGTCTCAACA AGCAAGATTC AAAAGCTGAC GTCTCCAGCA AGGCGGACTC GGAAGAGGAG 1560  
 CCGCTACTTC ACAACAGCAA GCTGGACAAA GTTCTGGGG TACAAGGGCA GCGCAGAGAA 1620  
 GAGGAAGCAG AGCGCTCTAA TCGAGAGCTC GCGTGTATGG GGAACAAGCA GCACCAAGGC 1680  
 AGGACAGATGG GGGCGGGGGC ACACCGCCCA ATGATCTTGC CCGTCCCTC CCGAAGAGC 1740  
 ACCACACTTA GCGAGTGGGA AGTGTTCATC CCGGAGCTGT GGAATACCAA CGTCTGCAG 1800

ACCCAAGAGC TSCGGCACTT CAATGCTCTT CTGGAAGGGA GCGAGAGGCC CCGAGCAGCC 1860  
 CCGGAGAGCG CTACGCTTTT CCGCAGGACCAA GAAGCGCAGC ATTCCGCCAA GGTCTCCACC 1920  
 AAGAGCCTCT CCAAGAAATG CCGTAGCCCA CCGTGGCGCG AGCGTGCCAT CCGTCGCGCA 1980  
 CTGAAGCAGA CCCCAGAGAA CAACCTTTGG CAGAGGCAGA AGAGGCTGCA GGCAATGCA 2040  
 AAGAGCGCCG TGCATGCTC AGTCTTTGA

Seq ID NO: 28 Protein sequence:

Protein Accession #: P123456 GENESh predicted

1 11 21 31 41 51  
 MSAGVAAGT RPPSPPTGS RRRQRPSVG VQSLRPQSPQ LRSQDPQKRN LDLEKSLQFL 60  
 QQQHSEMLAK LHEEIEHLKR ENKGPARGP RFALPPQARS TLPLQPHENT AINSSTRLSG 120  
 GGTQDGEPLQ TVLAHLAALA PVCQPSGYRF WGTWTDAA TS SRGWTMLCSQ AQHVLLSGSP 180  
 GPEVIAQRQV ATGCSPLPP PSRAEMGRNP WDSQPCARS L PQIAAVARPR ISSPMALSPH 240  
 MLGAGQWHT SIQSLPAIW AATMGTGKGS RVLPFCHLK ALPHPSGPH PAQDPQLWQ 300  
 AHFLSLZLG L TSGHLENG WSQGNIAAG A VPRALPSQ DMKEVSDGP FSRCKNSSE 360  
 LPWAKCPSPR QPQPCAGDA DRTREEMLS LGTCCSMCPK PSCFPDPSG NHLRSAPAL 420  
 GARWVCINGV VVPEQGPSPA RLKEGSSRTH RFGGKIRGLA GGSADTVRSP ADSLSMSSTF 480  
 SVKSSNSAN SQGKARQPG SPNKQDSKAD VSQKADLEE PLLHNSKLDK VFGVQQAARK 540  
 EKABASNAGA ACMGNQSHQG RQMGAAGHP MELPLPLPK TLRQCEVLI RELWNTNLLQ 600  
 TQLRLHLKS LBSGRQFQA A PEEASFPRDQ BATHEPKVST KSLSKCLSP PVAERAILPA 660  
 LKQTPKNPFA ERQKLQAMQ KRULHRSVL

Seq ID NO: 29 DNA sequence

Nucleic Acid Accession #: NM\_012319.2

Coding sequence: 138-2405 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 TCGTGCGCA ATTGCGGACG AGACCGCGTG TCGCGCGCTG GTAGAGAATT CTCGAAGACA 60  
 CAGTGGCGCT CTTGTGGAAC CTAACCTGG CGCGGTGGCG GCGCTGGGA CAACAGAGCC 120  
 CGGGAGACGA AGCGGCAATG GCGAGGAAGT TATCTGTAAT CTGTACTCTG ACCTTTGGCC 180  
 TCTCTGTAC AATTCCTCTT CATGAACATA AAGCAGCTCG TTCCCCCAG ACCACTGAGA 240  
 AATAATGTC GAATTGGGAA TCTGGCATT A TGTGACCTT GGCAATTTCC ACACGGCAAT 300  
 ATCATCTACA ACACCTTTTC TACCGCTATG GAGAAATAAT TCTTTGTCA GTGAAAGGT 360  
 TCAGAAATAA TCGTCTGAG ATAGAGTATA AGAATTCAT ATACACAT 420  
 ACCACGACCA TCATCTCAGC CACGAGCATC ACTCAGACCA TGAAGGTCA TCAGACCATG 480  
 AKCATCACT AGACACGAG CATCATCTTG ACCATGATCA TCACCTCTAC CATATCATG 540  
 TGCTGTCGG TAAATAAATG CGAAAAGCTC TTGCGCCAGA CCATGACTCA GATAGTTCAG 600  
 GTAAAGATCC TAGAACACG CAGGCGGAAA GAGCTCACCG ACCAGAACAT GCCAGTGTGA 660  
 GAAGGAATGT CACAGACATT GTTAGTGCTA GTGAAGTGAC CTCAACTGTG TACAACAGT 720  
 TCTCTGAAG AACCTCACTT CTAGAACA CA TAGAGCTCC AAGACCTGGA AACTCTTCG 780  
 CCAAAAGATG T AGCAGCTCC ACTCCACCCA GTGTACATC AAGAGCGCG GTGAGCGCGG 840  
 TGGTGGTAG GAAAAAATG GAATCTGTGA GTGAGCGCGG AAAAGGCTT ATGTATTCCA 900  
 GAAACACAAA TGAATACTCT CAGGAGTGTT TCAATGCATC AAAGCTACTG ACATCTCATG 960  
 GCATGGGCACT CCAAGTTTCG CTGAATGCAA CAGAGTTGCA CTATCTGTG CCAGCGATCA 1020  
 TCACCAAAAT TATGTCGTA TCTGTCTGTA TCTATACAG TGAAGAGAG GCTGAATCC 1080  
 CTCGAAGAGT TACTCATTA CAAATAGCCT GGTTGGTGG TTTTATAGCG ATTTCATCA 1140  
 TCAGTTTCTG GTCTCTGCTG GGGGTATCT TAGTGCTCT CATGAATGG GTGTGTTT 1200  
 AATTCTCTCT GAGTTTCTT GTGGCACTGG CCCTTGGGAC TTGAGTGTG GATGCTTTT 1260  
 TACACCTCTT TCCACATCT CATGCAAGTC ACCACATAG TCATAGCCAT GAAGAACCAG 1320  
 CAATGGAAT GAAAGAGGA CCACTTTTCA GTCATCTGTC TTCTCAAAAC ATAGAAGAA 1380  
 GTGCTCATTT TATCTCAGG TGAAGGGCTT TACAGCTCT AGAGGCGCTG TATTATGAT 1440  
 TCTGTGTA ACATGTCTCT ACATGATCA AACAAATTTA AGATAAGAA GCTGAAGAAT 1500  
 AGAAGAAACC TGAATAATGAT GATGATGTGG AGATATAGAA GCAATGTGTC AAGTATGAAT 1560  
 CTCACCTTTC AACAAATGAG GAGAAGATG ATACAGATGA TGAAGCTGAA GGCATTTTAC 1620  
 GAGCAGACTC ACAAGAGGCC TCCCACTTTG ATTCTCAGCA GCCTGCGAGTG TTGGAAGAAG 1680  
 AAGAGTGTAT GATAGCTCAT GCTCATCCAC AGGAGATGTA CATGTAATG GTTACCCAGAG 1740  
 GGTGCAAGAA TAAATGCCAT TCACATTTCC AGCATACAT CGGCGATGCA GACGATCTCA 1800  
 TTCACACCA TCATGACTAC CATCATATTC TCCATCATCA CCACCAACCA AACCCACCT 1860  
 CTCACATGCA CAGCGAGGCG TACTCTGGG AGAGGCTGAA AGATGCGCG GTGCGCACT 1920  
 TGGCTGTGAT GGTGATAAT GGTGATGCG TCACAAATTT CAGGATGAG CAGCAAGATG 1980  
 GTCTGCTTT TACTGTAAGG TTCAAGTG GTTAAAGTG TTGATGCTG GTGCTTTGG 2040  
 ATGAGTGGC TCATGATTA GTGCTTTTG GTTCTTACTT AAAGGTGGC ATGACCGTTA 2100  
 AGCAGCGTGT CCTTATAAT GCAATTGTG CAGTGTGCTG GTATCTTGA ATGGACACAG 2160  
 GAATTTTCAT TGGTCATTAT GCTGAAATG TTTCATGTG GATATTGCA CTATGCTGT 2220  
 GCTTATTCAT GTATGTGCT CTGTTGATA TGTGACTGTA AATGCTGCAC AATGATGCTA 2280  
 GTGACCATGG ATGTGACCG TGGGGGTATT TCTTTTACA GAATCTGTGG ATGCTTTGG 2340  
 GTTTTGAAI TATCTACTT ATTCTCAT TATGACATAA AATGTGTGTT GGTATAAAT 2400  
 TCGATTAAG GTTTAAATG TAGATAGCT TAAAAAGTTG TCATAGTTTC AGTAAATCAT 2460  
 AGGGAATGA GTTTGTATG TGTACTATG ACAGTTTAAA GTTAATGGGT TTGTGATTT 2520  
 TTGTATTGAA TATTCCTGT TGTACAAAG TCAATTAAG GTACUPTTIA ATATTATGT 2580  
 TATCTATCT TGGAGATAAA ATCTGTATG GCAATTCACC GTATATACCA GTTATATG 2640  
 TAAACAAGAG ATTTGGCATG ACATGTTCTG TATGTTTTCG GGAATAATGT CTTATAGCT 2700

TTTCAAGAA CTAACACAGT TATTCCTATA CTGGATTITA GGTCTCTGAA GAACCTCTGG 2760  
 TTTTAGAAA TAAGATGTTG CATGAAGCCT AAAATACCAA GAAGCTTAT ACTGAATTITA 2820  
 AGCAAAAGAA TAAAGGAGAA AAGAGAGAAGAA TCTGAGAAAT GGGAGGCAT AGATTCTTAT 2880  
 AAAAAATCAA AAATTTGGTG TAAATTAGAG GGGAGAAATT TAGAATTAG TATAAAAAAG 2940  
 5 CAGAAATTAGT ATAGAGTACA TTCATTAAAC ATTTTGTCA GATTATTTC CGCTAAAAAC 3000  
 GTAGTGAGCA CTCCTATATA CTAATTAGTG TACATTTAAC TTGTATAAT ACAGAAATCT 3060  
 AAATATATTI AATGAATTCA AGCAATTATC ACTTGACCAA GAAATTGGAA TTGCAAAATG 3120  
 TTGGGGGGG TATATAATCA CAGATGATCA GTAGATGATT TATGTATCAC CAGACTGGT 3180  
 TATTGCCAAG TTATATTATCA CCAAAGCTTG TATGACTGGA TGTCTGGTT ACCTGGTTTA 3240  
 10 CAAAATTATC AGAGTAGTAA AACTTTGATA TATATAGGGA TAITTAAACT ACACATAAGTA 3300  
 TCAATTGATT CGATTCAAGAA AGTACTTTGA TATCTCTCAG TGCTTCAGTG CTATCAATTG 3360  
 GAGCAATGTG CTTTATATAC GGTACTGTAG CCATACTAGG CCTGTCTGTG CGATTCTCTA 3420  
 GATGTTCTCT TTTTACACAA TAAATTCCTT ATATCAGCTT G

Seq ID NO: 30 Protein sequence:

Protein Accession #: NP\_036451.2

1 11 21 31 41 51  
 MARKLSVILI LTFALSVNTP LHEIKAAAP QTTEKISPWN ESGINVDLAI STROYHLQQL 60  
 FPKYGGNNSL SVEGFRKLQ NGIDKKRI HHHHHDHHS DHEHHSDHER HSDHEHSDH 120  
 20 HSHHSDHHS HEHHAASGN KRKALCPID SDSSGKDFRN SQKGGAHRPE HASGRNRVND 180  
 SVASSEVSTT VYNTVSGTH FLSTHETRP GLPLPKDVNS STPSVTSKS VRSLAGRKT 240  
 NESVSEPRKG EYNSAKSL ESHHSHHSH QNHRPHSSRGV VLSREELKA QVALTAWMV 400  
 25 RSLHTSEK KAEIPKTVS LQAVVGGFI AHSLSFLS LGVILVPLMN RVFFKLSF 360  
 LVVALVGTLS GDAFJLHLLP SHASHHSHS HEEPAEMEMKR GPLFSLHSQ NIEBAYPDS 420  
 TWKGILTALG LYFMLFVHV LTLKQFKDK KKKNKQKPN DDDVEKKQL SKYESQLSTN 480  
 BEKVTDDBKT EGYLADPSQE PSHTDSCQPA VLSEBVMIA HAHFQEVYNE VYPRGKNNK 540  
 HSHFDTLQG SDDLHREHD YSHLHSHHSH QNHRPHSSRGV VLSREELKA QVALTAWMV 600  
 30 MDGLRNSD GLAIGATTE GLSGLSTV AVPEHLHSH LGDFPVLKA GMTYQAVLY 660  
 NALSAMLAYL GMATYGFPIH YAEVNSMWIF ALTAGLFMYV ALVDMVPEML HNDASDRGCS 720  
 RWGYFFLQNA GMLLIGFML LISIPEHKV FRIN

Seq ID NO: 31 DNA sequence

Nucleic Acid Accession #: NM\_002184.1

Coding sequence: 256-3012 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 GAGCAGCCAA AAGCGCCCGG GAGTGCAGCT GGGCGGCCCG GGGCAGCTG AACCGGGGCG 60  
 CCGCGCTGCC AGGCGGAGCG GTCTGGGCCA GCTCGGGGCC AAGGGGTTG TGCTGCTGTG 120  
 40 GAGCGCGGAG GGTGCGAGCG GCGCGGCGCT AGTGAAACCC AATGGAAGAA GCATGACATT 180  
 TAGAAGTAGA AGACTTAGCT TCAAATCCCT ACTCCTTCAC TTACTAATT TTGTATTGGT 240  
 45 AAATATCCCG GCAAGATGTT GACCTTGCGAG ACTTGGTGAG TGCAGGCTT GTTTATTTC 300  
 CTCACCACTG AATCTACAGG TGAAGTCTTA GATCATGTG GTATATGAG TCTCGAATC 360  
 CCAGTCTCG AACTCTCTTA TATTTCATCT CGAGTTTGTG TGTCAAAGGA AAAAGTGATG 420  
 50 TATCTATTTC ATGTAAATGC TAATTACATT GTCTGGAAAA CAAACCAATT TAACTTCTCT 480  
 AAGGAGCAAT ATACTATCAT TGCAGAACCA GCATCCAGTG TCACCTTTAC AGATATAGCT 540  
 CATTAAATA TACAGCTCAC TTACCAACAT CTTCATCTG GACAGCTTGA ACAGAAATGT 600  
 55 TAGGAAATCA CAAATAATT AGGCTTGCCT CCAGAAAAAC CTAATAATT GAGTGCATAT 660  
 GTGACAGCTG GAGTGTGTGAG TGGGATGTG GAAGGGAAC ACACCTGGAG 720  
 60 AGCAACTTCA CTTTAAATC TGAATGGGCA ACACACAAGT TTGCATGTT CAAAGACAAA 780  
 CGTGACACCC CCACCTCATG CACGTGTGAT TATCTACTG TGTATTTGT CAACATGAAA 840  
 GTCTGGGTAG AAGCAGAGAA TGCCCTTGGG AAGGTTACAT CAGATCATAT CAATTGTGAT 900  
 65 CCTGTATATA AAGTGAAGCC CAATCGGCCA CATTAATTAT CAGTGTATCA CTAAGAGGAA 960  
 GTGTCTAGTA TCTTAAAT GATGCTGACC AACCCAGTA TTAGAGATGT TATAATACTA 1020  
 70 AAATATAACA TCTAATTAG GACCAAGAT GCTCTCACTT GGAGCCAGAT TCTCTCTGTA 1080  
 GACACAGCAT CCACCCGATC TCACTTCACT GTCCAAAGCC TTAACCTTT TACAGAAATAT 1140  
 GTGTTTAGGA TTGCTGTAT GAAGGAAGAT GGTAAAGGAT ACTGGAGTGA CTGGATGAAA 1200  
 75 TAGACCAAGT GGATCACCTA TGAAGATACA ACTCTAAAG CACCAAGTTT CTGTGTATAA 1260  
 GATATGTCAT CCCATATCA AGGCTACAGA ACTGTACAC TCTGTGGAA CTAAGTTCCT 1320  
 CTTTGTGAAG CCAATAGGAA AATCTGCTAT TATGATAGTA CTCACAGAS ATGGAAATCA 1380  
 CATTTACAAA ATTACACAGT TATGCGACA AACTGACAG TAAATCTCAC AAATGATCCG 1440  
 80 TATCTAGCAA CCCTAACAGT AAGAATCTT GTTGCGAAAT CAGATGCACG TTTTATAC 1500  
 65 ATCCCTGCTT GTGACTTTCA AGCTACTCAC CCTGTAAATG ATCTTAAAGC ATTCGCCAAA 1560  
 GATACCATCT TTTGGTGGGA ATGACACTAC CCAAGGGAAT CTGTAAGAAA ATATATACCT 1620  
 CAGTGGTGTG TTTATAGCA TAAAGACCC TTATACAGAG ACTGACACA AGAAGATGCT 1680  
 ACCGTGCATC GCACCTATT TAAAGGGGAC TTAGCAGAGA GCAAATGCTA TTGTATAACA 1740  
 70 GTTACTCCAG TATATGCTGA TGGACCAAGG AGCCCTGAAT CCATAAGGC ATACTTAAA 1800  
 CAAGCTCCAC CTTCCAGAGC ACCTACTGTT CGGACAAAAA AAGTAGGGAA AAACGAAGCT 1860  
 GTTCTAGAGT GGGACCAACT TCTGTGTUAT GTTCAGAAAT GATTATACAG AAATATACCT 1920  
 ATATTTTATA GAACCATCAT TGGAAATGA ACTGTGCTGA ATGAGTGTG TTCCACACA 1980  
 GAATATACAT TGTCTCTTT GACTAGTAC ACATTGTACA TGTACCAAT GGCACAGTAC 2040  
 75 ACAGATGAAG GTGGGGAAGG TGTGTCAGAA TCTACTTTA CTACCCCAA ATTTGTGCTA 2100  
 GGAAGAAATG AAGCCATAGT CGTGCCTGTT TGTCTAGCAT TCCATTGAC AACCTCTCTG 2160  
 GAGTGTCTGT TCTGCTTAA TAAAGGAGAC CTAATTAATA AACACATCTG GCTTAATGTT 2220  
 CCAGATCTCT CAAAGAGTCA TATTGCCAGG TGTGCACTCT ACACCTCCCA AAGGCACAA 2280

TTTAATTCAC AAGATCAAA GTATTCAGAT GGCAATTTCA CTGATGTAAG TGTGTGGAA 2340  
 ATAGAAGCAA ATGACAAAAA GCCTTTTCCA GAAGATCTGA AATCATTTGGA CCTGTTTCAA 2400  
 AAGGAAAAAA TTAATACTGA AGGACACAGC AGTGGTATTG GGGGGCTTC ATGCAATGCA 2460  
 TCTCTAGGC CAAGCATTC TAGCAGTGAT GAAATGAAT CTTCACAAAA CACTTCGAGC 2520  
 ACTGTCCAGT ATTCTACCGT GGTACACAGT GGCTACAGAC ACCAAGTTCT GTACAGTCAA 2580  
 GTCTCTCAAA GATCGACGCT TACCAAGCCG TTGTATGATT CAGAGAGAGC GGCAGAGAT 2640  
 CTCAAAATTAG TAGATCATGT ACATGAGCGT GATGGTATT TCGCCAGGCA ACAGTACTCT 2700  
 AAACAGAACT GCAGTCAGCA TGAATCCAGT CCAGATATT CACATTTTGA AAGGTCAAAG 2760  
 CAAGTTTCACT CAGTCAATGA GGAAGATTTT GTTAGACTTA AACAGCAGAT TTCAGATCAT 2820  
 ATTTCACAA CTGTGTGATC TGGGCAAAAT AAAATGTTTC AGGAAGTTCT TGCAGCAGAT 2880  
 CTTTGTGCT CAGTACTACT GGGCAAGTA GAAAGATTTT AAACAGTTTG CATGAGCT 2940  
 GGGAGTATG AAGATGCGT TGAATTTTAC TACCAAGTGA CTGATGCGCA AGGGGCTAC 3000  
 ATGGCTCAGT GAGAGCTAG TAGTTCCTGC TACAACITCA GCAGTACCTA TAAAGTAAAG 3060  
 CTAAAAATGAT TTATATGCTG AATTC

Seq ID NO: 32 Protein sequence  
 Protein Accession #: NP\_002175.1

1 11 21 31 41 51  
 | | | | |  
 MLTLQTVVQV ALFELFTBS TGEILLPCGY ISFESPVQL HSNFTAVCVL KEROMDYFHV 60  
 NANYVQKTN HTFPEQVQV IIRKTSVYV FTEIASLQV LTNCLTFQV LERQYVHT 120  
 ISGLPFEKFK NKSCVNEGR KMRCEWDGGR ETELENTFL KSEWATKIFA DCKAKRDTPT 180  
 SCTVDYSTYV FVNVEWVEA ENALGKVTSD HINFDVPYKV KPNPHNLSV INSEELSSL 240  
 KLTWTPNSIK SVILKYNYQ YRTKEDASTVS QIPREDTAT RSFTVQDLK PFTEYVPRIR 300  
 CMKEDGKGYW SDWSEASGI TYEDRPSKAP SFWKYDFESH TQGYRTVQLV WKTLPEFEAN 360  
 OKILDYEVTL TRWKSHE QNY TVNAKLTYN LTNDRYLATL TVRNLGKSD AA VLTIPACD 420  
 FQATPTVMDL KATKRNKGV VYKYLWVQLV SKAKATPQV QDEBQYVHRT 480  
 YLRGNLAESK CYLITPTPYV ADGPGSPESI KAYLKQAPFS KGTPTVTKKV GKNEAVLEWD 540  
 QLPVDVQNGF IRNYTTPYRT IEGNETAVNV DSHTEYTLS SLTSDILYMV RMAAYTDEGG 600  
 KDGPETFTTT PKFAGSEIBA IVVPVCLAFI LTTLLGVLEFC FNKRDLKIKH IWPNVDPESK 660  
 SHIAQWSPHT PFRHNPNKSD QMYSDGNFTD VSVVRIBED KFPFEDLKSL DLFPKKSEIN 720  
 TGEHSGGGG SSCMSSRRS ISSDDPERS QNTSTVQYS TVVSVYRIQV YHVVQVPSRS 780  
 ESTQPLLDHS ERPEDLQLVD IVDGDDGILP RQVYRQKCS QHSSSDSH FHSKQVSSV 840  
 NEEDFVRLQK QSDHSQSC GSGQMKMFEQ VSAADAFPGT TEGVQRUFET VGMEAATDEG 900  
 MPKSVLPQTV RGGGYMPQ

Seq ID NO: 33 DNA sequence  
 Nucleic Acid Accession #: NM\_018255.1  
 Coding sequence: 11-2491 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 | | | | |  
 AGTTGGGACG ATGTGGGAC CGTGCTGGA GACTTCTCAC GTGTTTGGCT GCCCAAACCG 60  
 GGTGGGGGGA GTCTGGAAC TGACCTCTGG GCACCAGAGA CTTCTGGCTC TTGGCACGCT 120  
 CTGCTCCGTG GTGCTCTATG ACCCCCTGAA AAGGGTGTGT GTTACCAACT TGAATGGTCA 180  
 CACCGGCGCA GTCAATTGCA TACAATGGAT TTGTAAACAG GATGGCTGCC CTCTACTGTA 240  
 ATTAGTTTCT GTGACATCTG ATAACTCAAGT GATTCACGTG GAAATAGAGG ATAACTCAGT 300  
 TTAAAGAGCA GTGCACTGTA AAGCCATGTA AGACATCTAT TATCGCGTCA ATCTGTCTTA 360  
 CCAGAGAGCA ACATCACTAT CTGATGATTA TACACTGATG GTTCTCTGCA CTGAGATTC 420  
 TGCCTTTCGA CTCGGTGCTA AAAAGGGTCC AGAAGTAATG TGCCTTCAGA CTTTAAACIT 480  
 TGGAAATGGA TTTCCTTGG CTTCTGCTTT ATCTTTTTCG CCAAACTACTG ATGTACCAAT 540  
 ATTAGCATGT TGGCAATGATG ATTGCAAGAT TTACATATTT GCTCAACAAA ATGATCAGTT 600  
 TCAGAAAGTG CTTCTTCTCT GTGGACATGA GGATTTGATT AGAGGAGTGG AATGGGACAC 660  
 CTTTGTGAGA GATCTTCTCT TAGCAAGCTG TTCAACAAGT TGCCGTATATA GANTATGGA 720  
 GCTGTATATA AAGTACACAT CTTTAGAAAC TCAGGATGAT GATACATAAA GACTGAAAGA 780  
 AAATACTTTT ACCATAGAAA ATGAAAGTGT TAAATATGCA TTTCCTGTTA CTTGGAGAC 840  
 AGTGCTAGCC GGTCAAGAAA ACTGGGTAAA TCGAGTTTCA TGGCAACCTG TGTTTACAA 900  
 AGATGTGTCT CTACAGACAGC CAGTGAAGAT TTATCTCTCT TCCATGGTAT AAACCATGAT 960  
 TCTGTGCTCT AGCTAGAGAG AGCTAGAGAT TTGCTAGAGA CAGGTTCCAG TAGGTGAAGT 1020  
 AGGTGGGAAT ACCTTGGGAT TTATGATTTG CCACTTCAAT GAAGATGGCT CCATGATCAT 1080  
 TGCTCATGCT TTCCACGGAG CGTTTCAACT TTGGAAACAG AATACAGTTA ACCCAAGAGA 1140  
 TGGGACTCCA GAGATTGTCA TTTCAGGACA CTTTGTATGT GTCCAAGACC TAGCTGGGA 1200  
 TCCAGAAGGA GAATTATTATA TCACTGTGTG TACTGATCAG ACAACTAGAC TTTTGTCTCC 1260  
 ATGGAAGAGA AAGAGCAATC CACAGGTGAC TTGGCATGAA ATTCGAAGGC CTGAGATACA 1320  
 TGGGTATGAC CTAAGATGTT TGGCAATGAT TAATCGGTTT CAGTTTGTAT CTGAGACAGA 1380  
 TGAAGAAAGT CTTCGGGTGT TTTCGCAACC TCGGAAATTT GTGGAANAAT TTGTGGCAT 1440  
 TACAGGACAA TCACTGAATC ATGTGCTCTG TAATCAAGAT AGTGATCTTC CAGAAGGAGC 1500  
 CACTGTGCCG GCAATGGGAT TATCAAAATA AGCTGCTCTT CAGGGAGATA TAGCTTTCTCA 1560  
 GCGTCTGAT GAAGAGGAGC TGTTAACATG TACTGTGTTT GAGTATCAGC AGTGTGGCTT 1620  
 TGAOCCCTCC ATACTCTGAT AGCCTCCGAC TGAGAGATGAT CTTCTGAGCA ATACTGTGTT 1680  
 CGCTGAAGTT CAATAACTAT ATGGGCCAGG TTATGAAATA TTTTGTGTTA CTGTGAACAC 1740  
 TTCAAGAACT CTGCTTGCTC CAGCTTGTTAA GGCAGCTAAG AAAGAGCAAT CAGTATCAT 1800  
 CTTTGTGAAC ACTACATCTT GGAAACAGGT GCAGAAATTA GTTTCCACA GTTGGACAGT 1860  
 CACGCAAGAT GCGTTCTCAC CTAATGAGAA GTTCTTACTA GCTGTTTCCA GAGATGCAAC 1920  
 TGTGTCATTG TGGAAAAAGC AGGATCAAT CTCACTGGAG TTGCAAGCCAG TTTTATGCTT 1980  
 TTTCGCTTCC ACCAACAAA TTACTCTGTG GCACAGTAGA ATATTGTGTT GTTGTGATTG 2040

GAGTCTGCAG ACGAAGTATT TCTTCACTGG GAGTGAGAC AAAAAGGTGG TTGCTGGGG 2100  
TGAGTGGCAG TCCACTGATG ACTGTATTGGA GCACAACATT GGCCCTGCT CTTCTAGCTCT 2150  
GGACGTGGGT GCGGCTGTGA CAGCTGTGAG CGTCTGCCCA GTGCTCACCC CTTCTCAACC 2200  
ATACGTGGTT GCAGTAGGAT TGGAGTGTGG AAAGATTTCG TTATATACCT GGA AAAAAGAC 2280  
TGATCAAGTT CCAGAAATAA ATGACTGTGAC CCACGTGTGA GAAACAAGTC AAAGCCAAAG 2340  
TCATCACTGG GCTATCAGAA AAT TATGCTTG GAAGAAATTC AGTGGAAAAA CTGACAGAGA 2400  
GGAAAGAGAA GGTCTGAT GTTACACACT TCCAAGCTGT GGTGAAGATC ACACGTGATA 2460  
GATACACAGA GTCAATAAAT GTGCACGTGA ATGG

Seq ID NO: 34 Protein sequence:  
Protein Accession #: NP\_060725.1

1 11 21 31 41 51  
MVAPVLETSV VFCCPNVRVG VLNWSSGPRG LLAGFTSCSV VLYDPLKRVV VTNLNGHTAR 60  
VNCIQWICKQ DGSFSTELVS GGSNDQVHW HIEDNQLLKA VHLQGEHPV YAVHAYVQR 120  
TSDPALCTLI VSAADSAYR LVBSKGFPHVM CLQTNFNGG FALALCLSL PNTDVPILAC 180  
GNDICRHRF AQNQNGQKVL LSLCCHEDWV RVVWFAAFR DFLFASQSD CLIRWKLKVI 240  
KSTLETQDD DNRILKENTF THENESVKIA FAVILETVLA GHENWVNAHV WQPVFYKDG 300  
QTPVVRLLSA SMDKTMLWA PDSESGVWLE QVRVGEVGGN TLGFYDQFN EDGSMIAHA 360  
FHGLRLHLWKQ NTVNPREWTF EIVISGHFDG VQDVLWDPGE EPIITVGTQD TIRLFAPWKR 420  
KDSQSVWTHIE IARFQIHGYD LKCLAMDNRF QVGSADSKV LRVFSAPRNF VINECATQQ 480  
SLRNVLNJD SLELZEGAV ALGLSNKAVP QGDIAQSPD EEELITSTCF EYQVVAQPS 540  
ILTEPTTEDI LQNTLWPEV QKLYGHGVEI FCVTNNSKT LLASAKAAK KEHAAAILWN 600  
TTSWKQVQNL VFHSLVTQM AFSPEKFKLL AVSRDRTWSL WKXQDTISP FEHFVSLEAF 660  
TNKITSVHSR IWSDCWSPD SKYFFTGSRD KKVYVWBCD STDDCIRHNI GPCSVSLDVG 720  
GAVTAVSVP VLHPSQRYV VAVLRCGKIC LYTWKIKTDQV PEINDWTHCV BLSQSQSHLT 780  
AIRLCKWNC SGKGTGKEAB GABWLHPASC GEDHTYKIRH VNKCAL

Seq ID NO: 35 DNA sequence  
Nucleic Acid Accession #: NM\_022131  
Coding sequence: 11-2878 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
TGCTGCGAGG ATGCTGCGCTG GCGCGCTGTG CIGGGTGCCG CTCTGTGCTG CGCTGGGGGT 60  
GGCGAGCGCG AGCGCGGGTG GCGGGGACAG CCGGACGCGC CGCTCTCTCG CGGCTAAAGT 120  
CAATAAGCAC AAGCAGTGGG TCAGAGACTTC ATATACATGGA GTTATAACTG AGAACAATGA 180  
CACAGTATT TGGAGTACG CCGTGATGCA AGATCATGCA GTTCACTTTTC CTTCTTTTC 240  
AGCGGAAATCT TGTGCGTATG AGATCATGCG CCAGGAGCTG CCTTGTGAGG CTGTGGTGTCT 300  
CAACAAGACA TCAGGAGAGG GCGCGCTCGC TGCCAAGAGC CCGATTGACT GTGAGTTGCA 360  
GAAGGAGTAC ACATCTATCA TCAGGCGCTA TGACTTGTGT GCTGGGCCCC ACAGACACAG 420  
CTGGA AAAAG TCACACAAGG CCGTGATCCA TATACAGTGG AAGGATGTGA ACGAGTTTCC 480  
TCCACACTTC AAAGACGAG CTTACAAGGC TGTGTGAGG GAGGCGAAGA TCTATGACAG 540  
CATTTCCAG GTCGAGGCC TGTACAGGA CTCTGCCCA CAGTACAGCC AGATCTGCA 600  
CTATGAATCT GTACACACAG ATGTGCCITT TGCCATGAC AGAAATGGCA ACATACGAGA 660  
CAGTGAAGG CTGAAGTATG ACAACAACA CAGATATGAG ATCTCGGTGA CGGCTACGA 720  
CTGTGAGAG AAGCGCCGCT CTCACGACAG CCGTGTGACG GTGGATGTGA AGCGAGTTTG 780  
CAAGCTGCG TCGCAAGACT GACCAAGAGI GATTGATGAC CAGCTGTGCT CGGCGAGCAT 840  
GCCCTGTC CCGACATGC ACTCGAGACG GTGCGA TGGGA GCCGTGTCTT CCTCCAGAT 900  
GTCGACAGG CTCGAGACTA ATTACATTGG GAAGGGTGTG GACGCGGAGA CCTACTCTGA 960  
GAAATCCCTT CAGAAATTAT GTGGAGCCTC CTTGGGTCAT ATTGACTCT TGCCATCCCC 1020  
TAGCGCTGCC ACCAACTGGA CTGAGGACTC CTCTGGTGAG AGCATCTGAGA TGTCTTCAA 1080  
GTTTGACGRC AGGCGAGGVS CC AAAATGCC CGATGGGATT GTGCCAAGA ACCTGACGGA 1140  
TGAGTTGACG ATACCATGAT GGATGAACA CCGCGCCAGC CTTGTGTGGA GAGCGACAGA 1200  
GGAACACATC CTCTGACAT CACTGAAATC CCAAAATGAC CGGATCATCAT ATGCGCTGTA 1260  
TGTCACAAAC TGCCGCTCTG TCTTTCTCTT GCGGAAGGAC TTGACACAGG CTGACACTCT 1320  
TGCCGCCGCG GAGTTCACAT GGAAGCTGGA CAGATATTGT GACAAAAGAT GGCACACTCA 1380  
TGTCATCAAT GTGAGATTTC CTGTGTAAAC TTGTATGAGG GATGAGACAA CATATGAACC 1440  
ATACCTGTGG ACCAAGCACT GGCCCATICA TCATATCACC ATAGCCATGC AATCATAGAT 1500  
CGCGCTGT TGCAAGAGAG GAGAAGTACG CCAAAATGAC TTGCTGATCTG TCTTCTA 1560  
AAGCTGCGCC AGTCTACCA TCGCCCTCTG CAAAATGGAA AGGCAAGAAG TGATCTCTGT 1620  
CTGTCAAGCC TGCAAGGAAG GCGTGGACAT TAATTTCTTG GAAAGCCTTG GCCAAGGAAT 1680  
AAAGATACAC TTCAACCCCT CCGATATCAT CCTGTGCTGG GAAGGTGAAG ACATTTGGAA 1740  
CATTAACCGT GCTCTCAGAA AAGTCTCTTA CATCACTCC AGGCAATGCC CAACGCGGGG 1800  
TGTGCGCCCG CTGAAAGAT CCTTCAA AAT CAGTGTCTT GGGGAGAGG TATCATCAG 1860  
TATCTCGTAG GTAGAAGCT ATGTGATGGT DCTCCAGGCC ATCGAGCCCC GGATACCCCT 1920  
CGGCGGCACA GACCACCTCT GGAGACCTCG TGCCCGATT GAAAGTGCCA GGGGAGTGAC 1980  
CCCTTCCCTT GATATCAAGA TTGTGAGCAC CTTGCGCAAA ACCGAAGCCC CGGGGAGCT 2040  
GAAACACAGA GACCCCAAT CAGAGTCTCT AGAGGA AATG CTTCATACAT TAGATTTCTG 2100  
TGCAATTTTG GTATGAGAGG GAGACTGTGA CCAAGGACAG GAGTGTCTGG AGCTCAACA 2160  
CACTGAGCTC GACCAAGAC ACTGTGAGCT CACTATTTCT ACTGACGGCT ACTCCATCTA 2220  
CGGTGTGGCG TCATATGAGC GCTATTGAGCA GTTGCTACAT CACATCCGCT ACCGCAAGCT 2280  
GGTCCCGGCT TCCCTTGAGG CCGCGGCTTT CCGGATTAAG TGCTCAGAAC TCAATGGGCG 2340  
CTACACTAGC AATGAGTTCA ACTTGAGGAT CAGCATCTCT CATGAGAGCC AAGTCTCAGA 2400  
TAAGGAGCAT GTCAATCAT TGATGTGCA GCTCCCTCT CTCAAGTCTC TCCATCATCT 2460  
TGAGTCCCGG AGTACATCC AGCAGATCT AGTGTGCCA AGCATGTGCA CAGTGTGAT 2520

CATCATCTCC GTGTGCATGC TTGTGTTTGT CGTGGCCATG GGTGTGTACC GGGTCGGAT 2580  
 CCGCCACCAG CACTTTCATCC AGGAGACTGA GCGTGGCCAG GAATCTGAGA TGGACTGGGA 2640  
 5 GAGTTCGTGC CTGACTATCA CAGTCAACCC CATGGAAGAA CATGAAGGAC CAGGGCATGG 2700  
 GGAAGATGAG ACTGAGGGAG AAGAGGAGGA AGAAGGCCAG GAAGAAGTGA GCTCCAGCAG 2760  
 TGGCTCTGAC GACAGCGAAG AGGAGGAGGA GGAAGAAAGG ATGCGCAGAG GCGAGACATG 2820  
 GCGAATATGA GCGAGTAAG GCAAGCTGGA GTGGATATGA TCCACCTCC CTTCTCAGTG 2880  
 CCGAAGGGTC TGCTTGCTGG CCGACATGTC CTTTGTGTAA ACCCTGACCC AGTGTATGCC 2940  
 CATGTCTATC ATACTCTACC TTGTGATGCT GTGACATGTC TGGGAAGGCC TTCTCAAGCT 3000  
 10 TCGTGGAGCC CACCCTTTAA GCCTTGGGCA CTCCCTGTGT TTCACTCCATG GGGAAAGTCC 3060  
 AAGAAGGCCA GCAATGCCAT CAGTGAAGAG TTCAGGATAG ACTTTGTGCT GTAGGCTCCA 3120  
 TTCTTGCCCT AAGTGTGCA GCATCCCTGAC TACGCTCTGT CAGAGTTTGC GTTGTGTTTT 3180  
 TCTTGCGGG AAGAAGGCC ACCCTTGTGT CACTGACTGC CCGAGCTTCA GAGTGTCCAA 3240  
 GGGCCTTGGG TTCCAATCA CTGTGCTGCT CCTCCACACA GACCACTAGG TTCTCTTATG 3300  
 15 CTGACTCCAG GTTGCTTCAT ACAAGGAGGG TGGTTGAATC TCACACACGT AAGGTCTTAG 3360  
 TTGTTAACAG TTAAAGGAAA AGTCTTGTT GAGCGACAGC TAAGTTTACA GGGAAAGGTA 3420  
 CACACATTTCT CTCTCTCTCT CTCTCTCTGT CTA TCTAGTGT CCGCAGCTTG GAGAGGCTTT 3480  
 CCGCTTGCTT TCTTCTAGG CCGATATAGC TTATAGAGAA AGTCCCAAC CAGAATTAGG 3540  
 TCTTGCGCA CAAAGCAGGT CTGATCCCCC ATCAGAGCTA TCTGAGCTGT CCGTGTCTGG 3600  
 20 CACCTGCTGC AACCATGACG CTACCTTGCC AGGGGCGACT AGCAACACAG ACCACAGGGC 3660  
 CCAGGAGGCA TTCCACACAG GCACGTGCCC AGGACAACAC AACAAGGACA CTCACACAAA 3720  
 GGACACAAAG GAACAACAC ACACACACAC AAGGACAGTC ACAACAAGCC TAGAGCCCAA 3780  
 AAGCAGATGG AATATCTAAT GAGTGTCAAC GTGAGCTTCA TGGTGGTGAG AGTGGGGTGG 3840  
 GCTGGGCTCC CCGAGGACAG AGGGGAGCCT GAGGTTGGCA AGGCTCTAC CACTCAAGCT 3900  
 25 TAGGTGCTCT TATCTCCTAT CTTCGCTCT GAGAAAATAC ACGCTTCTG CATGTATTAG 3960  
 AAACGCAQGA GCTCCACAAA GTCTACAATG AAGTGTGAAA ATTAACTGC AAGGAATAGT 4020  
 AAGCATATTT GCAATCTTGT CAGCTCTCTC TTCTCTCTGC TCATAAAGG AGGACAACCT 4080  
 TAGATAGAGG AACAATATAT CTGAAGAACT AATTCTTCTC TTTTTTGAT AAGGAATCT 4140  
 TTTCATCTTC CAACTCAAC CACTCTGTAAGAGAA GTTCTGTTAT AGTAACTGCT 4200  
 30 CTGTGATCTT TTGTGGCCAA GAGAAATGCA GCGCAAGATT ACGGCGCTGA CAGAAATTTCC 4260  
 ACQAACTCTC GAGAACATGT TTGTTTGAAA TGCTGTATTC CTCTTTGTCA TCAATGTGTC 4320  
 TGTCTGCTCC CCACTCTTCA CTCTCTCTCA AGCTCAACAC AATTGTTTGA GCACAGGCAC 4380  
 AGAGCTGTGT CTAAGTTAAG TGGCATTTAT GTTAAAGAAA AATAGTTCAG AATTCAGACC 4440  
 TTCTTCTGTG GTCTATCA GAGAGGACTA GAGAGGACTA AGCTCAATGC CTCTAGAGTC 4500  
 35 ACCTCAACC AGGGAGGAGC CATGGCAGGT CTCTTCAACT TTCTGATCA TGAGAACAC 4560  
 CTGTGTAAGC TTTTCCACC TCTAAAGTT TTCTGTGAT CTGTCTCTC CTCTTGACCT 4620  
 CACAACAAAT CCGTGAAGAT AACTGAGACA TCGTTTGTA GATACATTT TTGTGAGGT 4680  
 AAAGTGAAGC TTGG

Seq ID NO: 36 Protein sequence  
 Protein Accession #: NP\_071414.1

1 11 21 31 41 51  
 | | | | |  
 45 MLFRLCWVPI LLLALVQSGG SGGGDSRQR RLLAAKYNKH KPWTISYHGH VITENNDTVI 60  
 LPLVLALDK DAPVFPAGBI CAFKIRHOEL PFEAVVLNKT SGBGLRLAKS PIDCELOKEY 120  
 TFIQAYDCG AGPHETAWKK SHKAVYHIQV KDYNFAPFT KEPAYKAVVT BKIYDSILQ 180  
 VEAIDEDCSF QYSQICNVEI VTTDFPAID RGNRINRTEK LSYDKHQYB ILVATYDCGQ 240  
 KPAAGDTTLQ VDVKPKCKPG WQDWTKRIEY QFGSGSMPLF PSUILETCDG AVSLQVTE 300  
 LGTNVKGCC DREYSEKSL QKCGASRSH IDLLPSPA TNYVAGLLV SSMKPRK 360  
 50 RQCAKIPDGI VKNLIDTGT ITMYNKRKPS PGRVABKETI LCNSDKEMRN RHVIALVYNN 420  
 CRLVFLLRKD FDQADFTFPA EPHWKLQDQC DKRWYHYVIN YEPFVVLVYM LGATYEPYLV 480  
 TNDWPHPSH IMAQLTVYAC WQGGVEVTKP QAQFHGSLA SLTRFGKME SQKVSICLQA 540  
 CKBOLDINSL BSLGQGIKXV FNPQSILVM EGGDDGMNLR ALQKVSYTNS RQPTAGVYRR 600  
 65 LKYSKVKQCF GEDVYSIPE VDAYVNLVGA EKPRTLGGT DIEWRFAAQF ESARGVTLFP 660  
 DIEGVSTFAK TTAGVGYDCT DPKFVLRM LINDLPDREI VEGDLPREYR KLELINSHEL 720  
 55 HQRHLDATNS TAGYSYIVGQ SMSRYEQLVH HIRYKYNRPA SI EARFRPRK CSBELNGRYTS 780  
 NEFNLEVSIL HEDQVSDKEH VKNLIVQFP LQSVIHPEER SSQHSVSVF SIATVVIHS 840  
 VCMLEFVVMAM GVYKVRIRHQ HFIQTETAAK ESEMDWDISA LTTIVNPMK HEGPHGHEDE 900  
 TGEHEEEAE EEMSSSGSD DSEEBEBEG MGRGRHGQNG ARQAQLEWDD STLPY

Seq ID NO: 37 DNA sequence  
 Nucleic Acid Accession #: none found  
 Coding sequence: 143-874 (underlined sequences correspond to start and stop codons)

65 1 11 21 31 41 51  
 | | | | |  
 GGGAGGAGAGA GAGGCGCGCG GGTGAAAGGC CGATTGATGC AGCTGCGGCG GGCCTCGGAG 60  
 CCGCGGCGGAG CCGAGAGCGCT ACCACGTTCC TCTCTCGGT CTCTCCGCG TCCAGTCCG 120  
 70 GCTGCGCCGG CAGCGCGGAG CCGTCGAGAC CCGAGGCCCC GCGCGCTCC CCGACGCGCT 180  
 CCGCGGCGCT CTGCTGCTCC TGTCTGCTGA CCGCGCCCG CCGCGCGAGG CCGTGAAGAT 240  
 CCCCAGGGG AAGTCAAG CCGAGCTCCG CCGAGAGGAG GTGCTGCGACG TTGATTAATGG 300  
 AATTGTCCTA CAGGCGGAG CAGGAGTCCG TGCTCGAGAC GGGAGCCCTG GGGCCAAATGG 360  
 CATTCCGGGT ACACCTGGGA TCCAGGTGCG GGAATGAATC AAAGGAGAAA AGGGGGAATG 420  
 75 TCTGAAGGAA AGCTTTGAGG AGTCTCGAC ACCCAACTAC AAGCAGTGT CATGGAAGT 480  
 ATTGAATAT GGCATAGATC TTGGGAAAT TCGCGAGTGT ACATTACAA AGATGCGTTC 540  
 AATAAGTCT CTAAGATGT TTGTCAGTG CTAACCTGG CTAAATGCA GAAATGATG 600  
 CTGTACGCTT TGGTATTCA CATTCAATGG AGCTGAATGT TCAAGACCTC TTCCATTGA 660

- AGCTATAATT TATTTGGACC AAGGAAGCCC TGAATGAAT TCAACAATTA ATATTCAATG 720  
 CACTCTCTCT GTGGAAGAAC TTGTGGAAGG AATTTGTGCT GGAATGATGG ATTTGTGTAAT 780  
 CTGTGTGGGC ACTTTGTCAG ATTACCCAAA AGGAGATGCT TCTATGGAT GGAATTCAGT 840  
 TCTCTGCATC ATTTATGAAG AACTACCCAAA ATAAATGCTT TAAATTTTCA TTGCTACCTC 900  
 TTTTITTTAT ATGCTCTTGA ATGTTTCACT TAAATGACAT TTTAAATAAG TTTATGTATA 960  
 CATCTGAATG AAAAGCAAAG CTAAATATGT TTACAGACCA AAGTGTGAT TTCAACATGTT 1020  
 TTAAATCTA GCAATATTTA TTGTGCTTCA ATCAAAAAGT GTTCTCAATAT TTTTITTAAG 1080  
 TGGTATAGAT ACTTATCTCA TAGTCAAT TTCTCAACCT ATAAATTTGA ATATTGTGTG 1140  
 GGTCTTTGTT TTTTCTCTCT AGTATAGCAT TTTTAAAAAA ATATAAAGC TACCACATCT 1200  
 TGTACATTT GTAAATGTTA AGAATTTTT TATATCTGT TAAATAAAAA TTAATTCCAA 1260  
 CAACCTTAAA AAAAAA AAAA
- Seq ID NO: 38 Protein sequence  
 Protein Accession #: none found
- 1 11 21 31 41 51  
 MRPGTAAAT QRLRLGLLLL LQLPAPSA SEHPKQKQKA QLRQREVVDL YNGMCLQGPA 60  
 GVRGRDSFGF ANGIPTGPI PGRDGFKEH GECRLRSFSE SWTPYKQCS WSSLYNIDL 120  
 GKIAECTFTK MRSNRLKVL FQSLRLKCR NACQQRWYFT FPGABCSGLP PHEAIFYLDQ 180  
 GSPDMNSTIN HRTSSVEGL CEGIGAGLVD VAIVWGTCSL YPKGDSTGW NSVSRILIEE 240  
 LPK
- Seq ID NO: 39 DNA sequence  
 Nucleic Acid Accession #: NM\_000949  
 Coding sequence: 285-2153 (underlined sequences correspond to start and stop codons)
- 1 11 21 31 41 51  
 GGAGGCTGAA ATCCCCAGAC GCGGTTTTC TGGGCTGGGC TTCTGCTTCA CTCATCTCTT 60  
 CTCCTCTCTT CTGGATTTTA CGAGCCGTTT CGGAAGAAGT TTCCACACA ATGGAGCTTC 120  
 ATGTCTCTGT CGAGGAAGTA CTCATCGACT GATGTGGCAG ACITTGCTCC CTGACAATAAC 180  
 TAAAGAACTC TCTATTCAT GGAGGGGCAAC ACTGAGGATG CTTCACAT GAACCTGTGA 240  
 GTGAACCTCT GATACATTC CTGAGCAAG AGCAAGGCA GC CAACGAAG GAAAAATGTG 300  
 CATCTCCAC ACCTTTTCACT CTGGTACTTT TTCTCAAGAC CTGCTCTCTG AATGGACAGT 360  
 TACTCTCTGG AAAACCTGAG ATCTTTAAAT GTCTTTCTCC CAATAAGGAA ACATTCACTT 420  
 GCTGTGGAG GCTCTGGGCA GATGGAGGAC TTCTACCAA TTATTCAGT ACTTACACA 480  
 GGGGAAGGGA GACACTCATG CATGAATGTC CAGACTACAT AACCGTGGGC CCCAATCTCT 540  
 GGCATTTTGG CAAGCAGTAC ACCTCCATGT GGAGGACATA CATCATGATG GTCAATGCCA 600  
 CTAAACCAT GGGAAAGCAT TTCTGGGATG AACTTTATGT GGACTGTACT TACATGATTC 660  
 AGCCAGACCC TCTTTGGAG CTGGCTGTGG AAGTAAAAAA GCCAAGAAGC AGAAAACCCCT 720  
 ACTGTGGAT TAAATGTCT CCACCTACCC TGATTGACTT AAAAACTGGT TGGTTCACGC 780  
 TCTGTATGA AATTGTGATTA AAACCCGAGA AAGCAGCTGA GTGGGAGATC CATTTTGTGT 840  
 GCGACGAAC AGAGTGTAAAG ATTTCTCAGC TACATCCAGG ACAGAAATAC CTGTCTCAGG 900  
 TTCCCTGCAA ACCAGACCAT GGATCTCTGA GTCCATGGAG TCCAGGAC CTTCATTCAGA 960  
 TACTATGTA CTTCACCATG AATGATACAA CCGTGTGGAT CTCTGTGCTT GTCTTCTGTG 1020  
 CTCTACTGT TTTGATTA TTCTGGGCGAG TGGCTTTGAA GGGCTATAGC ATGTGTACCT 1080  
 GCATCTTTC GCCATCTCT GTGCGCAAAA TAAAGGATG TGAATGCTAT CTGTGGAGA 1140  
 AGGGCACTCT TGAAGAACTA CTGATGTGCT TGGAGTGCCA AGACTTCTCT CCCACTGTCT 1200  
 ACTATGAGGA CTGTGCTGTG GAGTATTTAG AATGTAGATG TAAATGAGCA CAGCATCTAA 1260  
 TGTCACTGA TTCAAGAA CACCCAACTC AGGTATGAA ACCACATAC CTGATGTGCT 1320  
 ACATGACTCT AGGCGGGGAG AGCTGTGACA GCGCTTCTCT TTTGTGAA AAGTGTGAGG 1380  
 AACCCGAGGC CAATCCCTCC ACATCTCTAT ATCTGAGGT CATTTAGAGG CCAGAGATC 1440  
 CTGAACAAC CCACACCTGG GACCCCGAGT GATAAAGAT GGAAGGCAAA ATCCCTTATT 1500  
 TTGATGCTGT TGAATCCAAA TGTTCACAT GGCCTTTAC AGACCCGAGC CAGCAACAAC 1560  
 CGAGTCTCT TTACACAAAT ATTACTGAT GTGTGAGCT GCGTGTGGGC CCGTCAAGTG 1620  
 CACCGGCAC TGTGTGTAAT GAAGCAGGTA AGATGTCTTT AAAATCTCTT CAACACCTTA 1680  
 AGTCTAGAGA AGAGGGGAAA GCAACCCAG AGAGGGAGGT AGAAAGCTTC CATTTGAGA 1740  
 CTGACAGGA TACGGCTGTG CTGCTGCCCC AGGAGAAAAC CCGCTTGGC TCGGCTAAC 1800  
 CCGTGTGATTA TGTGGAAT AT CACAAGGTCA ACAAGAATGT TGCATATCA TTCTACCAA 1860  
 AACGAGAGA GACACAGGCG AGCCCAAGA AGCCCGGAG TCTGAGAAC AATAAGGAT 1920  
 ATGCCAAGT GTCCGCGTCA TGTGATGAA CATCTGCT GTTGTGGCA GATCCACAGA 2040  
 CTAAAAAGT GTCTTGTCTT GAAGAATCA CCAAGAGGC CCCAATCA CTGGAACAGA 2100  
 ATCAAGGTGA GAAAGCCCTG GCCAACTTCA CTGCAACATC AAGCAATGC AGGCTCCAGC 2160  
 TGGGTGTGTT TGATTACCTG GATCCCGCAT GTTTTACACA CTCTTTTAC TGAATGTTG 2220  
 ACTAAATGAA TGAATGGTTA AAATGTGATT TTCTCTCAGG TACATATCA GAGTADTTGA 2280  
 AATGCTCAG AATGTATCTA GATGTCACT ACTAAAGTC CCAAGTCTT TCAATGCTCA 2340  
 TTTTAAACA CTTCGCTCTT TCTCCAGCAG CTGATTCAG AACCAATCAT TATGTTTCT 2400  
 AACTGTGATT TGTAGATTTA CTTTTGTCT TTAGTTATA AACATATGT TCAATGAAT 2460  
 AAAAGCACAC TGTCTAGTCT TCTTGAGGA CAATGCCAAT AGGTATATCT TCTGGAATAAG 2520  
 GCTTTCATGA TGTCTGATGG GACAGACGGA AATGAATG TCAAAATGT TTTACATAGA 2580  
 AAGATGACAA AAGTAAAT TCCAATAGG AATAGGCAAT GAAATTTGCT TTTGAAAAAC 2640  
 AACTGATAA CTCTTACACT CCTCTCCCAT TTTATAGGA TTACCCAAAT ATAAACATT 2700  
 AAAGAAGAA TGAATCCAG AACCAATGT TTACATAGT TCCATACCT TACTGACACA 2760  
 TTGCTGATAT GCAAGTAAAGA AAT
- Seq ID NO: 40 Protein sequence



Protein Accession #: NP\_000940.1

1 11 21 31 41 51  
 5 MKENVASATV FTLLPLNTLC LLNGQLPGK PEIFKCRSPN KETFTCWWRP GTDGGLPNTY 60  
 SLVYERDGT LMRHEDYTH QVNSCHQK QVTEBWRVTV MMVATNTQMGS SPSDELYVD 120  
 VTVVYQDPD LELAVEVQK EDRKPYLVWK WSPFTLIDLK TGWFILLYE RLKPEKAAEW 180  
 EHFAGQQTE FKILSLHPGQ KYLVQVRCKP DRHYSWAWSP AITQIFSDF TMNDTTVWIS 240  
 VAVLSAVICL IYVWAVALKG YSMVTCFPP VPGPKIKGFD AHLEKKGSS ELLSALGQD 300  
 10 FPFSTDYEDL LVEYLVDSDS EDQHLMVSHV KEHPSQGMKP TYLDPDITDG RQSCDSFSL 360  
 SEKCEHFQAN PSTYDEYEV EKPENETH TWDFQICIME GKPYFHAGQ SKCSTWLPQ 420  
 PSQHNPSYV HNTDVCILA VGPAGAPATL LNEAGKDLK SSQTKREB GKATQREVB 480  
 FSHSETDQDT PWLLPQEKTP FGSAKLDDY EBHVKNKDGA LSLPKQREN SGKPKKGTTP 540  
 15 ENNKYAKVS GVMDDNMLVL VPDHAKNVA CFEEBACEAP PSLEQNQAER ALNFAATSS 600  
 KCRILQGLLD YLDPACFTHS FH

Seq ID NO: 41 DNA sequence

Nucleic Acid Accession #: none found, Bos cloned sequence

Coding sequence: 1-1572 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 20 ATGACCCAAA ATAAATTAAA GCTTTGTTCC AAAGCCAATG TGTATACTGA AGTGCTGAT 60  
 GGAGGATGGG GCTGGGGCGGT AGCTGTITCA TTTTCTTCG TTGAAGTCTT CACCTACGGC 120  
 25 ATCATCAAGA CATTGTGTGT CTCTTTTAAT GACITAAATG ACAGTTTAA TGAATCAAT 180  
 AGCAGGATCT CATGGATAAT CTCAATCTGT GTTGTGTCTT TAACATTTC AGCCGCCCTC 240  
 GGCAAGTCC TGAGCAATCG TTTGGACAC GCTCTGGTGA TGATGTGGG GGGGCTACTT 300  
 GTCACACCG GATGCTGCG CCGCTCTCTC TCACAGAGAG GTTCTCTAT GTACCTCGC 360  
 30 ATCGGCATCA TCTCTGCTT GGGATCTGCG TTAGTITTC TCCCACTGT AACCATCTA 420  
 TCACAATATT TTGGCAAAAG ACGTTCATCA GTCACTGCGA TTGCTCCAC AGGAGAATGT 480  
 TTGCGTGTGT TGTCTTCCG ACCAGCAATC ATGGCTCTGA AGGAGCGCAT TGGCTGGAGA 540  
 TACAGCCTCC TCTTGTTGGG CCTACTACAG TTAACATGT TCACTCTGG AGCACTGCTC 600  
 AGACCATCT TTATCAAGG ACCAGGCTCA CCGAAATAG TCATCCAGGA AATTCGGAAA 660  
 GAAGGCGAGT ATATCTTGA AAATAGAGAA ACACGAACCT CAATGACTC CATTGACTCA 720  
 35 GGAGTAGAAC TAACATCCTC ACCTAAAAAT GTGCTACTC ACCTAAACCT GGAACCTGGG 780  
 CCGAAGGGCG ACATGCAGCA GGTCCTGGTG AAGAACAGCG CCAGGCCAAG CGAAAGAGAA 840  
 CGCCGCGTAT TAGACTTCTC CATTITGAAA GAGAAAAGTT TTAATTGTTA TGCATATT 900  
 GGTGCTTTTG CAACTCGG ATCTTTTGA CTTCTCTGT ACATATCTC TCTGGCAT 960  
 40 AGTCTGSCA TTGACAGGA CGCGCTGCT TTTTATTAT CTACGATGGC CATTGCAGAA 1020  
 GTTTTGCGAA GGATCGGAGC TGGTTTGTG TCACAAAGGG AGCCATTGC TAAGATTATC 1080  
 ATTGAGCTCA TCTGCGTCT CTTATTGACT GTGCTCTGT TTGCTTTAC TTTTGCTACT 1140  
 GAATCTGGG GTCTAATGTC ATGCAGCATA TTTTGGGTG TATGGTTGG AACAAATAGA 1200  
 GGGACTCACA TTCACTGCT TGTGAGGAT GATGTGTGG GCATGAGAAA GATGCTCTT 1260  
 45 GCACTCGGG TCTACTCTT CATTGAGAG ATAGCAGGAC TGGCTGGACC GCGCTTTGCA 1320  
 GGTTTGTGG TGGACCAAG TAAGATCTAC AGCAGGGCCT TCTACTCTG CGCAGCTGCG 1380  
 ATGGCCCTGG CTGCTGTGT CTTGCGCTG GTGAGACCGT GTAAGATGGG ACTGTGCCAG 1440  
 CATCATCACT CAGGTGAAC AAAGTATGT AGCCATCTG GGAAGACTT ACAGGACATA 1500  
 50 CTTGAAGACT TCTGGAATG GGATCTTGA AAAAATGAGC ACAGAGTTCA CTGCAAAATG 1560  
 GAGCGGATAGA

Seq ID NO: 42 Protein sequence:

Protein Accession #: none found, Bos cloned sequence

1 11 21 31 41 51  
 55 MTQNKKLCS KANVYTEVPD GGWGWAVAVS FFFVEVFTYG IKFTGVFFN DLMDSFNESN 60  
 SRISWISIC VFVLTFAPL ATVLSNRFH RLVMVLGGL VSTGMVAASF SQBVSHMYVA 120  
 IQISILGYC PSFLPTVTL SQYFGKRRI VVAVASTGEC FAVFAFAPAI MALKHRIQWR 180  
 60 YSLPLVGLIQ LNIIVFGALL NFPIFGPAS PRIVQBNKK EAPYMLBNK TRISIDIS 240  
 GVLETFBN VFTFENLEB PKADMQVLY KTSFPRSEK AFLPLSILK BKSFYIALP 300  
 GLFATLGFPA PSLYILPLGI SLIGDQRAA FLSTMAIAE VFGRIGAGPV LNRBFIRKY 360  
 45 HELICVILLT VSLFAFTAT EFWGLMSCI FPGMVVTGQ GTHIFLLAE DVVGIEKMSS 420  
 AAGVYIFQSI AGLAGPLA GLLVDQSKY SRAPYSCAA MALAAVCLAL VRPCKMGLQC 480  
 75 HHSHOETKVY SHRGKTLQDI PEDFLEMDLA KNEHRVHYQM EPV

Seq ID NO: 43 DNA sequence

Nucleic Acid Accession #: FGENESH predicted ORF

Coding sequence: 1-1749 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 70 ATGCTGTCGT GCTTCTGAT GAGTCCAGT ACCGACGACA GAGCACAGTA CACTCCGGA 60  
 GGAAGAAAC TCCGTGGGA GGCTTCATC GTTGGGACA CTTCCGGAGG GCGAGGCAGC 120  
 75 GACCGGGAGA GGGAGAGCG CCGCAGAGCT GCGCGGCTCT TGTGGGACG CCGCTGACGC 180  
 GGGAGAGCGG AGAAGGGGAA CCAGGCGCAG CCGCCGCGCT GATAGCGCG CCAGACBAG 240  
 CCGCGGCGC GCGAGCGG GCGAGCGCG GCGAGCGCG CTGGGGGCG CCGAGACCGT 300  
 CGCTGCGTC CTGACCTTC CCGGGGAGG GTCCGTTGC CAGTGAACC TCCAGAGCT 360

TCCGAGCAGC AGCCCCGGGG GCCTTCTGAC TCATCCCGA GATTTCATC AGCGAGTGCA 420  
 ATCTATAAGC CAGTCCCTAA GGGGACGCGG CCACCGGGCT AGGACGGGGA TGGCTTAGGA 480  
 GCTCTTGAC CTAGGGCCCG GGGTGTGTGG CTCTCTGGCG TGGCGGCGA GGGGAGTGCG 540  
 CGCGCGGAAA AGCGCGCGG GACAGTGAAG TACGAGGCGC GGGGGTCGCG GGGGCGACGA 600  
 CTCTGTGGAG ACGTGGCTG GCTCTGTGGA GACGCGCTGT CGGCGCCGAG GGTGGTGGCA 660  
 TGTGGGGCGC TGGCGCTGCT TCGTCTCTCT CATCTCGGAA CGCGCGCTCG CTCTCGAGC 720  
 TGTCTGTGCG TGGCTGTGCT GGGGGCGGGG CGAGGGGCCA GCGGGGAGTA CTGCCACGGC 780  
 TGTCTGGAGC CUCAGGGGCT CTGGCGCATC GGCTTCCAGT GTCCCGAGCG CTTCACGGC 840  
 GGGGACGCGA CCAATCTTGG CGGCACTGCG GCGTGTCTCT ACTGCTGCTC CAGCGCCGAG 900  
 GCGCGCGTGA ACCGCGGCT GCTGGAGATC GACCGCGGAC AGGGCGGCGC GGGGCGGCG 960  
 CGGCGCGGCA AAGACGGGCG CCGACGCGCT GCGAGGCGCT CATGTCTTAG GGTATCCCAA 1020  
 GGAGACGGCG AGGGTGGCG CCCACCCTGG AGGGGCTGGC AGGGGTGCTC CCGTGAAGCG 1080  
 TCCCGGAAA GAGGCGAGCT CCTCAGGGCT TTCCCGGGCG TGCTGCCCGC TGGCAGACGC 1140  
 CGGGAATTC CATCTTCTCC ACBGGCGGCG CCTCTCCGCC TGCAGCGGGCC CGGCTTGCCC 1200  
 ATCTAGTGC GTTCTCTCAT TGTGTGCTCC GTTGTGTGCT GCTTATCAT CTGGGGTCCC 1260  
 TGTGTGCGAG CTTGTGTGCT CAGATGTCTC CGGCTTAAGC AGGATCGGCA CAGGACGCGA 1320  
 GCGCCAGGGG GTAAACGCTT GATGGAGAAC ATCCCATATG TCACCAGTGC CAGCACTCTC 1380  
 CGGGGTCTGT CCTCAGCGCA GTCCAGCACA GCTGCCAGTT CAGGCTCCAG GCGCACTCTC 1440  
 GGGGCGCGGG CGCCGCCAAC AAGGTCCAGC ACCCACTGTT GCITGCCGGA AGGGACCATG 1500  
 AACCAAGTGT ATGTCAACAT GCCCAGCAAT TTCTCTGTGC TGAAGTGTCA ACCCGGCGAC 1560  
 CAGATGTGC CACATCAGAG GCGAGTATCT GATCCCGCAT ACGTGGGTTA CAGGGTGTAG 1620  
 CAGCAGTCTG TGGCCATGAC AGCTGTGCGA CTTTTCATGG ACGGCTGTGA GCTCGGTCTC 1680  
 AGGCAGATTC AGTCCCTCTT CCTCACAAGC AACAGTGAAC AGAAGATGTA ACCCGGCGTG 1740  
 ATCTGTAAAC CGAGATGTCAT TGGTGGGTTCT CTTTACTGAA GGGGAGCGAA GCGAGCGGGT 1800  
 GATCTCGAG GTGGAGGTCC GCACATGTGCG GTGTATTTTA TGGACAGATT CTTTGGTAGT 1860  
 GGTTCATTTG CCGGCGGAG GTATGTAAAC ATCTCGCATC TACGATCTTT GATATGTTT 1920  
 CATCCAGGTT ATCATATGAT TATGATGGAA AACCGGCGCT ACBTGGAGAT GATGTGTGAT 1980  
 TTGCTGATGG GTGTATAACA AATGCTTGAG TCCGAAGTGC CTTGAGATA TGGTGTACGA 2040  
 AAGAATTTTA TAAACTGATA AATTAAGGAT TTTTATTATG TTGTTATTAT TATTTCTTT 2100  
 TTGTTGTGA CTGACAGGA TCAAAATGCC TGTATCTCC CTTTACTGCG GACTCTTTTT 2160  
 TTTTCTTTT TTTTCTTA TCAAGCAGGG TGTGTGCTGT TGGCCAGCG TGAATGTGAG 2220  
 TGTGTGATC TGGGCTGCT GCAACTTCAG CTTCTCGBAT TACGAGCACA CTTCTGCTC 2280  
 AGGCTCCAC CTTGGCTGGGA TTACAGGTGC CTGCGCCCAT GGCTAATTT TGTATTTT 2340  
 TGTAGAGATG GGGTGTCAAC ATGTTGGCTG GGCTGGTCTC ACTCTCTGTA CCTCAAGCAA 2400  
 TGTGCTGTCT TCAGCTCCCG AAAGTGTGTG GTATTACAGG GTGAGGCCAC GCGGCCAGCG 2460  
 TTAGGCTTTT TTTTCTCTA ATGCATCCAA GTTTAAGGGG AAGACGCGAA TAAAGAGGAT 2520  
 ATTCATAAG GAAATCTGTT GAGATGATCT ATCAGTCTGA GTATTCAGA 2580  
 GCGACACCTT AATTCTATG TAAAAAGATA TATATATTTT GTCTATTTT GTGCTTTTGG 2640  
 GGGCGTATTT TGTGCTTTT TACCATTATG AGAGATCTTA TTACAAGATG ATTCTTACA 2700  
 TTTAAAAGAG ACTGAATAAA ATTGTATAGT TACTTAAC TAAGAAGACAT TCGAGAAGT 2760  
 TGGGATGATT TTAATCTTGA AGTAGTAGGT GGTATAGTCA TAAACCAATT CATCCCGCTT 2820  
 TGTATGTAT CTAATATCT TGGCTTTAAG GTGACATGCG AGAGGTAAAT GACTCTTTTT 2880  
 TATATTGAAA TCAATAACTA TCAACGCTGT CTTCTCTGAG TTAAGTTTAA TTTTGCTGT 2940  
 TGTATTGTGT TTGGGTTTCT CTCTGTTTG GTTTTCAGAG CCCCATTGCT ATATAGTCT 3000  
 GAGTGCAGT AATTACTATA CTGTAAATG AAGATCAGTA TTTCGCTTA GATGTGATA 3060  
 AAAAAATTTT TGTCTTAGT TATAAAAAAT CAAAGAAATG TGTTCAGAA GATCTTAGTA 3120  
 TAACTCTCA CCAATAACCT GAGACTTGGG ATGAAAAATTA AACCGATAC GATTACTTTT 3180  
 CGAGATCATA AGGCTTTT TACTGTGTT ATCAAAATGG CTTATTTTC AGGCACTAAG 3240  
 GATTGTAAAG AGAAAGCTT TCAACGAAG GATGGCTTTT CTTCTCCAC ACTGTTCTG 3300  
 ATTTCCTCT CTCTTCAGCG CTCACAGCGC ACTGTATTCA TTGCGAATGT TCCAAATTAT 3360  
 CAAATTCAGG TGAATTTATT TGTGTGTTCT TTAATTAT TAAAAAAGAT AACCTAAGG 3420  
 ATTGCGAAGT ACATTTCCAA CTGTACGAC AACCAAGATT TTGTATTTAA AACTATCGCT 3480  
 GTATGATG GTATGAGA CATTTATGTC TATAGATATC TATGCAATCT CTTTCTATC 3540  
 TTTTCTATG A CTGAATATG TAAACCAACT GTTGGCAAT GTGTATCA ATGATATCTA 3600  
 TTTTATAATA ACCAAAGGCA GGGGAAATC ATTTTACTTA TTAATAAATA TTTTATGATG 3660  
 TGAATAAAAA AAAAAAATA AAAAAAATA

Seq ID NO: 44 Protein sequence  
 Nucleotide Accession #: P0ENSH predicted

1 11 21 31 41 51  
 | | | | |  
 MLSGFLMSPS TQHRAYTFG QKFLVWBSI GAHTSRGRGS DRERESRPEA AGLLWDRAAA 60  
 GEABGRNRGE PPAVRAQQ RRPAPCAQAP GTAAAGCAQF RLKRGSRGK VRLPVPEPA 120  
 SGROPGRPSD CPAPFESA BATHAVPKGT PPAEDDGLGL AFGPKARRRR LLGVAAEGSS 180  
 PRGKRRGTVS DBAROSPGR LLGDRPALSP DALSAPRVVP GCAALAPRSP HPGTPLRSCS 240  
 CCLWLRCWRIG RPSGGEYCHG WLDAQGVWRI GQCPERFGD GDATCCGSC ALRYCCSAE 300  
 ARLDQGGCIN DRQAGAGHP RADKDGPRLR GRASCLRGTF GDGEGAPPPV RAWQRCSBPG 360  
 SFKRGQLLRA RPLLRARR RQPPSSRPGG RPSLQRLALP TVVFLVGS VYVAFILGS 420  
 LVAAECRCLL RPKQDQQRH AFGNRLMET PMPISASTS RGSRRSST ABSSSIANS 480  
 GAKAPFTRSQ TNCCLPRTM NNVYNMPTN FSVLNCQQT QIVPHGGQYL HPFVVGVTYQ 540  
 HDSVMTAVP PFMGLDQGY RQIQSPFPH NSQKLMYPAV TV

Seq ID NO: 45 DNA sequence  
 Nucleotide Accession #: NM\_002285  
 Coding sequence: 55-3738 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 GGCGAGCCT CGGCGCGGCG GTGAGCGGCG GCGCGACGCG TGACACCTCC CACCATGGAC 60  
 5 AGCTTCGACT TAGCCCTGCT CCAGGAATGG GACCTCGAGT CACTGTGTGT CTATGACGCA 120  
 GATAGLAAT GATACCGAG G AAGAGAACGA GAAGAAGAGAA ATCAAGAACTCA ACAAGAGT 180  
 GATGGCACTT TTAATCTAG TTACTCTCTC TTCACTGAGC CTTACAGAC TAACAAGGGG 240  
 GATGAACCTT CCAACCGGAT CCAGAGAACT TTAGGCAATT ATGATGAAAT GAAAGACTTT 300  
 10 TTAACGTGTA GAACCAATCA GAGTCATCTC GTTGCGAGTC CCAAACTGGG GTTCTCTCAG 360  
 ACTCTGTGTA ACAAGATGTA TGAACATTTT GTTGCAGATT CAAGAGGCGA GAGCAGGCC 420  
 TCGTCTATCT GTAGCACTAC AACTTCCACA CCACAGAGCTG TCCCGTGCA CCAAGAGTAA 480  
 AGAGGCTA TGGGTCTGTA GAAGGCTGGG CACCCACCTT GTGAGGCGCA ACAGAGAGCA 540  
 ACACAACAAG GCTCTCTCAG GACCTTGCTT GGAGATGGTG TTGGCAGACA GCAGCTCCGG 600  
 GCGAAACAAG TGTGCAATGT GGAGTGGGCG CTTACAGCC AGGAGAGGCC ACCTGCCATG 660  
 15 GCGGCCAAGC ACAGCAAGCAG CGGACACTGT GTTCAGAACT TTCTCCATC CCTAGCTTCA 720  
 AAAAAACAAG TGGTTCAGCA GAACCGACCC GGTATGTGTA GCGCAATGTA CGGCCAAGAT 780  
 CAGCGCCCTC ATGAGTCTCC TACTCTGAAK TGGTCTTGG AAACAGAGCT GAGCTGACGA 840  
 TCATACAGGG GATCTCCCTG CAGCAAGCGG GAGCTGTGCA GAGCGAAGCG CAAGCTCTCC 900  
 AAGTTGACGA TCCCAAGCA GGGGGAGGAG AGTAGATCTG GAGAAACCAA CAGCTGTGTT 960  
 20 GAAGAATAAA TCGGGAGAT GACCTGGCTT CCACCACTT CTGCTATTCA AGCACTCGGC 1020  
 AAAGTGGAC CAACCAAAAT TCCATTGCA AATAAGGACT CTGAGCTTGT ATCTCTGGA 1080  
 CACAATAAT CAGTGGTGA GAGTTCAGAG CCAGAGAGTC CAGCAAAATG CATCATGAA 1140  
 ACATCATGTC TGAAGAATGA CTTTAAGCTA AGCAGTGTGA AAGAGAGAA TGAACGAG 1200  
 GCAGCTCAGA GAACGCTCT CGCGCTCTC TCTGACAGCG CGTGGTGCA GCAGCCCAAC 1260  
 25 TGCAAGAACT CGTGCTCTC CAGCAAGGCC AGCAGCAGCA GCAGCAGCAG CGGCAAGGCG 1320  
 AGCTCTCTCA GCGACTCAGA GAGCAAGTCC GGATCTGACT CGGAGACCGA GAGCAAGTCC 1380  
 GCGAGAGAGT AGGGGACGAA GCGCCCGCAC TTCTCCAGCC CCGAGAGCTGA ACCGGCATCT 1440  
 TCTAACAAAT GCGTCTCTA AACAATAAT TATCCCAACA GCGCTCTAT 1500  
 CTGTACTCAA ATGAAGAAGCA CGGGTCCAGG AGCAATAGCT ACTCAACCC GTGAAAGAG 1560  
 30 GAQTTCCAGG ACTGTGTGGAA AGTCCCGCAG GTTTCGCA CGCAGCTGAG AGAGAAGGAG 1620  
 ATCAAGAGCA CTTCAAGGA GAGCAAAAGG CCAAGGACAG CCAACAAGGC CCGTGGGAGT 1680  
 AAGAGCGTGA AGCAGAAGTC CCGCGCCGCG GCGCTGGGCG TGGCGGTGAG CCGCAGCGCC 1740  
 CGGCAACCGC CAAGTGGCGCG GAGACAGCGC CGCGCGCTCC CCGGAGGTCC 1800  
 35 GCGGGCAAGA AGCCCAACCB GCGCACCGAG AGCACTCAG CCGGGACAGG CGCCAACTCG 1860  
 CACCGCGCCG AGGAGCCGCG GCGCGCGGAC GCGCTGGGGA CGAGCTGTGT GTTCCCGCCG 1920  
 GAGCCACCA AAAAAAGGCC CTGTGGCAAC AAGAGAGGGA GCGACGCGAA GAGCTCGGCG 1980  
 TTCTCCGTGA CTTGGGAGAA GCGCGCGCAC CGGGGGCTTA GCAAGGATCT CCGCAAAATCC 2040  
 AAGAGGTGTA TTGAGACAGA GTGTGTCTAT TCACTCTCTC CTCTGCACTC GCGACTGAG 2100  
 40 TCGAGCAG AGGATGAGC TGTGTCTCAA GCACAGAGCC TGGTCTGCTC TGGCTCTCT 2160  
 GGAATGATCT AGAGGCTGAA GAGGCGCCTG GCCAAGCGGG GCAGTGTGTC TAGGGCCDCT 2220  
 TTAGGCTCCA TCAAGGCCAG GACCAACCAAT GACATGCGCA AGGAGCTGGA GGAGCGCTCT 2280  
 TACACACTGG TCCCTCTTGG CGGGAAGGAA CTCTCTCC CCCTAAAGGA CAGTGAAGAG 2340  
 45 ATCAGTCTCT TGTGGTTCAA AATCGACTGT ACCCTCTCTCT CCAAGATGTC AGAAGCACTG 2400  
 CCGCAGGAGC CAGGCGTATT GAGCGCCDCT GCGACCAAGG ACTCTGAGAG CGCACCGCC 2460  
 AGCCACACTC CCGACACACC TCGAAGAAAG GCTTTGCCAA ATCCAAAGG GAAACCGCAAG 2520  
 TTGTACAAGG AAGAGACTTA CAGGGAGATC AAGAAGTCCC AGGGAGAGAA AGACAGCTCT 2580  
 TCAAGACTGG CCACCTCCAC CAGTAATACT TTGTCTGCAA ACCACTGCAA CATGACACTC 2640  
 AACAGTGTGG CAATACAAT AAATAAAAAA GAAAAAATGC TTGGTGGCG CATCTCACCC 2700  
 50 CTCTCTGATG CACTTAAACA CAATACACC AGGAGAGACT TAACCTCTTC CAGCGCACT 2760  
 AATGTGCA GATTTTTATC TTCAGCTCTT TCCAGAAATA AGCTTAAGGCG CGACAGGCG 2820  
 CTGACGCTC ACGGCGGCA CACTCAGAAA GCAAGCTACA ACAATTCTGA AAACATTCOC 2880  
 CTCACAAAT CAGGCGCGCA GACGAAGCGG TGGTCTCCAG GCTCCAAAGC CCACAAGGAG 2940  
 TGCAGAGGCG AGAAACTTGT CTTGATGAT ATGCTCGCA GTGCGGATTA TTTATGCAA 3000  
 55 GAGACTAAAC GAATGAGCA TAAAGAGAT GCAATGTGTG AAAAGTTTGG AAAGGCTTTG 3060  
 AACTATGCTT AAGAGCA TT GTCTTTATC GATGTGTGAA ATGCAATGGA ADAAGGCC 3120  
 ATGAAATCCA AATCTCTCTA TTACTCTATG TATTGAGAAA CAGTAGAGCT CATCAGTAT 3180  
 GCTATTGAGC TAAAAACCCA CTCAGGCCCC AATGCCACAG CAGAAGACAA ACAACTTGCT 3240  
 GCAATTATTT ACCGATCCT GCGCCTCTG TACTGGGGGA GTTTTGACT CAAAAGGGAG 3300  
 60 CAGCTGTAAA AGTATTCAAA AGCACTAATC GACTATTICA AGAATCTATC TAAAGCGGCG 3360  
 CAGCGCCAT CAGGCTGAGA AAGAGACTGA GAGACCAATC CCGCATTTCT 3420  
 CCGCAACDCT TTCCCGGAG CTTCCGTGGG TCTCAGGCGA GCTCTTCAA CCGCAGCGCT 3480  
 CTTGTCGCTG GACCATGCT CAGCATGCGA CAGCGATCC ACCAGATGGC GGCCAACCA 3540  
 GTGACATCA CCAACAGCAT CCTGCACAGC TACGACTACT GGGAGATGGC CGACAACCTG 3600  
 65 GCGAAGGAAA ACGGAGAAAT TCTCAAGCAG CTCGGATCTG TCAATGGGCG GTTCACCGCT 3660  
 CACAGAGCAA TGGAGACTCT GTTCCAGAT TCCACAGAG GCTCTGAGAG GCTCTGGAG 3720  
 AGCGCAAGC TGTGTAAGG AACTCTAGCT GGTGCCAGAG TGGGTCTGG TGTCCACAGA 3780  
 TGGCTCAAGC TTTTGTGACA CTGTGCTACT GAAACTTCCA GCGACAGCAT TTATGAGCT 3840  
 CGGTGAACAT TTCTCA

Seq ID NO: 46 Protein sequence:  
 Protein Accession #: NP\_002276

1 11 21 31 41 51  
 75 MDSPLALLQ EIWDLISLCVY EPDRNALRRK ERHRRNQETQ QDDGTNFSY SLSPFQPKYN 60  
 KQDELNRN NQKTRTQGS HLVGTVPRGV FQTPNKNIDE IPVADSRAGN 120  
 QPSSCTTT STPAAPVQK SKRGTMGWCK AGHPRPDQO RATQQSLEK LGGDVGKQ 180  
 PRAKVQNVN VGLQTQRRP AMAAKHGGG HCVNQFPFL ASKPSLVQK PTATYVRMDG 240

QDQAPDPSKPK LKSSSETSVH CTSYRGVPAS KPPEPRAKAK LSKYSPKQGG EBSRSGETNS 300  
 CVBEHREBT WLPJLSAQK FGKVEFTKPP PPKNDISLV SHGNPNPKGD AEPSPDNGT 320  
 NNTSMLEDDL KLSDEEENE QQAQRTALR ALSDSAVVQK PNCRTSVPS KGSSESSSSG 340  
 TSSSSSDSS SSGSDSETES SSSBSBGSKP PHFSPRAEP ASSNKWLQJK WLNKNVNPFKP 360  
 PILQNESHG SHNQYNNPV KEDVQDCGKV PVOQPSLRE KEIKSTCKIE QRPRTANKAP 380  
 GSKDKVQKSP PAAVAIVASA AAPPFAVPCA PASNAPAPAR KSAGKPKTR TFKTSAGDGA 400  
 NCHIEPEEA ADALGTSVYV PPPTKTRPC GNWRASHKKE LRSVYCIEX KTRGLSRIVP 420  
 KSKKHETES SSSSSSDSD LSEGEQEPYL SKAQTVASA SGNDRQIKE AAANGSGGFR 720  
 APVGSNART TSDIAKELEB QYTLVPFR NELLSPKDS DEIRSLWVKI DLTLLSRIFE 780  
 HLQPEQGLVS APATKDSESA PPSHTSDTPA EKALPKSKRK RKCNDENBYR EIKKSQGEKD 840  
 SSSLRTATST NTLGASHNCNM NINSVAIPH KNEKMLRSP SPLSDASKJR YTSIDLTSSS 900  
 RPNONSIFTS ASSSKPKPAD SQLQPGQJDL TKAARHNSN IPIHKSRPT KFWBZGK 960  
 RDCRQKLPV DDMPTATYV MQEAGDMERK ADAMVEKFGK ALVNAEALS FIEGNGAMEQ 1020  
 GPMEKSPYV LMVSEITVELI RYAMRLKTHS OPNATPEQDK LAALCYRCLA LLYWRMFRLK 1080  
 RDHAKVYSKA LIDYTKNSK AAQATSPWGA SGEKSTGTPS ISNPNFPGSS VGSQSLNA 1140  
 SALSPSTVS IPQRHQMMA NHVSITNSIL HSYDYWEMAD NLAKENREFF NDLDLLMGVP 1200  
 TLHSSMHLV QYQQQLHLWL RNSAHS

Seq ID NO: 47 DNA sequence  
 Nucleic Acid Accession #: NM\_003151  
 Coding sequence: 351-449 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 ACTGGGATAA AGCAAGAAGA CTGATTTTAT GAGCAGGGGT TTGATACATC AAAGGAGATT 60  
 GCGGCAAGT AAGGTGCGG TGTTGGGGGT GGGTTGGGGA GGGTGGTTAG AGAAGGTTTC 120  
 ACTAAGTAT TTGGGCTGTA GCGCTGAGAA GATGTTTAAA AAGAGGGATC AAGCAAGC 180  
 TAAGGAGAG ATGCTGATG GACCTCAAG CTTGCAATGG GCCAATATG TCCTCTGACG 240  
 GTAGTGCCCT TTCTCTGGC TGCTCAAGC GAGATCTAAG CTCTTCTTAA CTCTCTGCT 300  
 CTTTTCATAT TCTCTGATC TGGGAAGCA AGAATTGGCA GGAACGTGAA ATGACTAGGA 360  
 AGAGGACATA CTGGGTGCC AACCTCTCT GTGGCTGGT GAATGTGGC ATGCACTAG 420  
 GCGATGACAT GGTTCAGGA GTTATTATA AAGCTATAC TCCTCAAGAT GCGCCCTGGA 480  
 GTACGACAGA GAGAAATCT GAGCTCCAG GAGGGGACG TGTGCAAGC TGGGGGAAGT 540  
 ATGATGCTGC CTGAGGAAC ATGATTCCT TGGTCCCAA GCGAGGTTT CTGCTGCCCC 600  
 AGCCCTCGGA CAATGCTGC CTGTTCTCT ACCTCAAGC GTCACTGGCT ACCCCGCTCA 660  
 TATCCAAAG CTTACGGAGT CGCTTAGATG AGAACACCAT CCTCTCACTG TCAGTCCATG 720  
 ATGCTCAGA CAATAATGTC CAAAGGCTTC ACXGCTTTG GGAAGAAGAA GTCTCAAGCG 780  
 GAGGGAATGA AAAGAGGAG AGAGATGAG TGAATGCTG GTTCCAGCA ACAAGTTGGA 840  
 TTTTGATGC ACTCTGCGC ATCTGCTCT GATTCGCAG TACTCTGGG CCAATATTGA 900  
 TTATACCAA GATCTGGAA TATTCAAGAG AGCAATGGG GAATGTGTC CATGGAGTGG 960  
 GACTGCTCT TGCCCTTTT CTCTCGAAT GTGTGAAGTC TCTGAGTTC TCCTCCAGTT 1020  
 GGATCATCAA CCAACGACCA GCCATCAGT TCGAGCAGC GTTTCCTCT TTGCTCTTTG 1080  
 AGAAGCTCAT CCAATTAAAG TCTGTATAG ACATCACTC AGGAGAGGCT ATCACTGCTT 1140  
 TCACCGGGA TGTAAACTAC CTGTTGAAG GGGTGTGTGA TGGACCTCTA GTACTGATCA 1200  
 CTTGGCCTC GCTGTGCTAC TGCAGCATTT CTCTCATCT CATATTGGA TACACTGATC 1260  
 TTATTGCCAT CTATGTGCT CTCTGCTTT TCCACTGCG GGTATTCAAT ACAAGATGG 1320  
 CTGTGAAGCT TCAGCATCAT ACATCACTGAG TCACGACCA GCGCATCCGT GTGACCAATG 1380  
 AAGTCTCAC TTGCTTAAAG CTGATTAATA TGTACATCAT GGAAGAACCA TTGCAAAA 1440  
 TGAATTAGA CTTAAAGAG AGACTTGA AACTTATGA GAATGTCGG CTGTGCTAGA 1500  
 GCGTACAGAT TATAACTCT TTCACTATC CCACATGGC CACAGCGCTC TGGGTCTTCA 1560  
 TCCACATCAT CTAAAGCTG AAACCTCAG GTCAATGGC CTTCAGCATG CTGGCTCTCT 1620  
 TGAATCTCT TGGCTGTCA GTGTCTTT TGTCTATGC ATGCAAGGT CTCACGAATT 1680  
 CCAAGTCTG AGTGTAGAG TTCAAGAAAT TTCTTCCA GGAAGGCGCT GTTTCATG 1740  
 TCAGACATTT ACAAGACCTC AGCAAGGCT TGTCTTGA GAGGCGCAC TTTCATGCG 1800  
 AACAGATCT TCCCGGCTC GTCAATGGC CACTGTGAGT GGAGAGGAAC GGCATGCTT 1860  
 CTGAGGGGAT GACCAAGGCT AGAGATGACC TGGGCGAGA GGAAGAAGG AACAGCGCTG 1920  
 GCGCAAGAT GCACAAGAT AACCTGGTG TGTCCAAGG GATGATGTTA GGGGTCTCG 1980  
 GCAACAGGG GAGTGTGAAG AGCAGCGCTG TGTCAAGCAT CTTGAGGAG ATGCACCTG 2040  
 TCGAGGCTCT GGTGGGGGT CAGGGAAGCG TGGCTATGTT GAGAGGAG GCTGTGATCT 2100  
 TCAGCGGAA CATACGGAG ACATCTCTA TGGAGGCGC ATATGACAG GCGCATATCG 2160  
 TCGAGGCTCT CCACGTCTC TCCCTGAAT GGCACCTGGA ACTCTGCC TTGGAGABA 2220  
 TGACAGAGAT TGGAGAGCG GCGCTCAAC CTCTGCGGG GCGAAGAAG AGGATCAGCC 2280  
 TGGCCCGCG CGTCTATCC GACGTGAGA TCACTGCTGT GGAAGCAACC CTGCTGCTG 2340  
 TGGACGCCA CGTGGGGAAG CACATTTTG AGGATGTCAT TAAGAAGACA CTCAGGGGA 2400  
 AGAGGCTCT CCGTGTACC CACAGCTGC AGTACTAGA ATTTGTGCG CAGATGAT 2460  
 TGTGTGAAA TGGGAAAAT TGTGAAAAT GAACTCACG TGAGTAAAT CAGAAAAAG 2520  
 GGAATATTC CCAACTTAT CAGAAGATG ACAAGGAAG CACTTGGAG ATGTGTCGG 2580  
 ACACAGCAA GATAGCAGAG AAGCCAAAG TAGAAAGTCA GGCTCTGGC ACCTCCCTG 2640  
 AAGATCTCT CACAGGAAT GCTGTGCGG AGCATACCT CACACAGAG GAGGAGATG 2700  
 AAGAAGCTCT CTGAGTTTG AGGGTCAAC ACCACTACAT CAGGACAG GAGGTGATC 2760  
 TGTGCTCTG CATAAGTCT TCTTGCTGG TGTGATGCT CTCTTAAAG ATCTTCACT 2820  
 TCTGTGCTG GAGTACTTG TTGAGCAGG GCTCGGGGAC CAATGACAG CSAGAGAGCA 2880  
 ATGGAAACAT GGCAGACCTT GGCAACATTG CAGACAAAT TCAACTGCT TTCTACCAAG 2940  
 TGGTGTAGG GCTCAAGCG CTGCTCTCA TCTGTGTGG GGTCTGCTCT TCAGGATTT 3000  
 TCCACAAAT CACGAGGAAG GCATCCAGG CCTTCACAA CAAAGCTCTT AACAAGGTT 3060  
 TCGCTGCCC CATGAGTTT TTGACACCA TCCCAATAG CCGGCTTTG AACTGCTG 3120

CAGGGGACTT GGAAACAGCTG GACCAGCTCT TGCCCATCTT TTACAGAGCAG TTCTGGTCC 3180  
 TGTCTTAAT GGTGATCGCC GTCTGTGTGA TTGTCAGTGT GCTGTCCCA TATATCTGT 3240  
 TAATGGGAGC CATAATCATG GTTATTGGCT TCAITTTATTA TATGATGTTT AAGAAGGCCA 3300  
 TGGTGTGTT CAAGAGCTGT GAGAACTATA GCGGTCCTCT TTATTCTCC CACATCTCTCA 3360  
 ATTCTCTGA AGCTGAGG TCACTCATG TCTATTGGA AACTGAAGAC TTGATCAGCG 3420  
 AGTTTAAGAG GCTGACTGAT GCGGAAATA ACTAAGCTCT GTTGTGTCTA TCTTCACAC 3480  
 GATGSAATGCG ATTGAGGCTG GAGATCATGA CCAACCTTGT GACCTTGGCT GTTGCCCTGT 3540  
 TGTGGGCTTT TGGCATTTCC TCCACCCCTT ACTCCTTTAA AGTCATGGCT GTCAACATG 3600  
 TGGCAGCTT GCGGTCCAGC TTCCAGGCCA CTGCCCGGAT TGGCTTGGAG ACAGAGGCAC 3660  
 AGTTCAAGCG TTGABAGAGG ATACTGCACT ACATGAAGAT GTGTGTCTCG GAAGCTCTCT 3720  
 TACAGATGGA AGGCGAGT GGTCTCCAGG GGTGGCCACA GCA TGGGAA ATCATATTT 3780  
 AGGATATCA CATGAAATAC AGAGACAACA CACCACCCTT GCCTCACGCG ATCAACCTGA 3840  
 CCATCCGCGG CCACGAAGTG GTGGGCATCG TGGGAAGGAC GGGCTCTGGG AAGTCTCTCT 3900  
 TGGGCATGGC TCCTCTCCGC CTGGTGGAGC CCATGGCAGG CCGGATTCTT ATTGACGGCG 3960  
 TGGACATTTG CAGCATGGC CTGAGGAGCT TGGCTGTCAA GCTCTCAAGT ATCCCTCAAG 4020  
 ATCAAGTCTT GCTCTCAGA ACATCAGAT TCAACTAG AGA TCCCTTTCAG CUTCACACTG 4080  
 ACCAGCAGAT CTGGGATGCC TTGGAGAGGA CATCTCTGAC CAAAGCCATC TCAAAGTTC 4140  
 CCAAAAGCT GCATACAGAT GTGGTGGAAA ACGGTGGAAA CTCTCTGTG GGGGAGAGGC 4200  
 AGCTGCTCTG CATTCGCCAG GCTGTGCTTC GCAACTCCAA GATCATCTT ATCGATGAAG 4260  
 CCACAGCCTC CATTGACATG GAGACAGACA CCGTGATCCA GCGCAAACT CGTGAAGCTC 4320  
 TCCAGGGCTG CACCGTGCTC GTCAATGCCC ACGGTTCAC CACTGTGCTG AACGTGTGAC 4380  
 ACATCTCTGT TATGGCAAT GGGAGAGTGG TAGAATTTTA TGGGCGGAG GTACTCGGGA 4440  
 AGAAGCTGG GTCACTTTTC CGAGCCCTCA TGGCCACAGC CACTTCTCA CTGAGATAAG 4500  
 GAGATGTGGA GACTTCATGG AGGCTGGCAG CTGAGCTCAG AGGTTACAC AGGTGCGAGT 4560  
 TCGAGGCCCA CAGTCTGCGA CCTCTCTGT TGGAGATGAG AACCTTCTCT GGAAGCAGAG 4620  
 TGAATATGTG GGGGTGTGG GATTCTGGA TGGAAACCTT GGAATAGGCT ACTGTATGCG 4680  
 TCTCAAGACC TTACAGTCTT AGACATGACT AAGACATGAG ATTCAAGTGT CATGTGTGT 4740  
 CTTCTTTAAT TACATGCTG AATAATTITA TAATAAGCTA AAAGCTTATA GTTCTGTACT 4800  
 CTGTGTAGA AGTGTGCAA ATCTGTACT GACTTGTATA AATATAAAAC TAAGAGAAAC 4860  
 TC

Seq ID NO: 48 Protein sequence:

Protein Accession #: NP\_149163.2

MTTRKRTYWPVNSSGGLVNRIGDIDMVSGLIKYTYTLQDGPWSQERNPEAGRAAVPPWKGKYDAALRT  
 MIPERPKPRPAPQLNDAGLSYLYTSWSLPLMIQSLRSKRLDENTPLFVSHDASDKNVQRLHRLWEEB  
 VSKRIEIKASVLLMGRQGTQLFDALLGICFASVLGPIILPKILEYSEELQNVVHYGLCPAF  
 LSECVSLSSSWINDXCTAIRFAVSSAPBRLIKPFSVHTTSBLSBEFTGTVYVFEQYCPY  
 VLTICASLVICSSSYFIQYTAFAILCYLLVPELFAVMTIRMAVKAQHITSEVSQRRIRVTSVLTKIK  
 LRMYYTWBKPFIKIEDLRKERKLEKGLQVSLTSTLPIIPTATAVVYLHTSLKLLKLTASMAFSM  
 LASNLRLSLVFFVPIAVKGLTNSKSAVMRFKPFILQBSVPVYVQLQPSKALVFEELTSWQQTCPG  
 VNGALEKNGHASIGMTFRDALGPBEBGNLSGPBLIKNLVYSKMMMLQVCGTGTGKSSLSAILEB  
 MHELEGSYVQCLAYVYQCAWYNSONIRENIMLGGAYDKARYLQVHLCCSLNDELHLPDMTEIGR  
 GLNLSGQKQRISLARAVYSDRQVLYLDPLSAVDHIVGKIDFECCKTLRGKTVTVLTHQLYLEFG  
 QHLLNGKICENGTHSELMOQKGYAQIQQMKHEATSDMLQDTAKIA3KFKVESQALTSLESLNLS  
 AVPEHQLQEBEEMBEGLSWRVYHYHQAAGGYMVSCITFFVVLVLTIFSWWLSYVLBQSGTNSL  
 RESNGTMDLGINADNPQLSFYQVYGLNALLICVQVCSGIFTKVRKASTANHLKNLVFRCPMSF  
 FDTTPGRLNCFAGLEQLDQLPFSBQFVLISLAIVLILVSVLYLLMGAIMVICTRYNMF  
 KKAJVPVKKLWYBSLPIYSLNQLSSBYHSGFQSLDQSNLYLFTSSTWAMFL  
 EIMTGLVTLVAALFVAFGISSTPSKFMVAVNVLQALSSQFATARGLEBTAQFVAVERILQYMKMVS  
 EAPLHMEGTSCTQGWQFHIEQDYHMKYRDNTPVTLHGNI.TRGHBVVGIVGTGSKGSSGLMALFR  
 LPMAGMRILIDGDCISGLDRLSKSVIPQDPVLVSGTIRFNLPDDRHDIDQWDALERTFLTAI  
 SKPKKLLHTVDYENGNGNFSVGERQLCIARVLNRNSKIKIDEATASIMEDTDLTQRTIREAFQCGTVL  
 VIAHRYTVTLNCDHILYMGKRVYFDRPBLIKKPKSLPALMATATSSLR

Seq ID NO: 49 DNA sequence

Nucleic Acid Accession #: NM\_033419

Coding sequence: 18-980 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CGAGCCAGG AGAAAGGATG GCGCGCCTGG CCGCGCGGTT GGTCTGCTCA GCTGGGCGAG 60  
 CCGCGCTCG GAGCGCGCTCC CAGGGCGGACC GTGAGCCGGT GTACCCGGAC GGTCTACTCG 120  
 ATGTGCAAGA CAGAAGCTGC TCTGGGGGCG CTTGAATCA CTCCTCCCTC CGCGACGCAA 180  
 TGTACATGAG TATGATGAGT GTGAGCACTG TAAATGATGAG TGTATGTGG 240  
 TCACCGCTGG GCTCTACTCT CAGGAAGGTC ACAAAAGTGG TCAAGTTCAT GCGCAATGGC 300  
 CTTCTCCGCG GTTCTGTTTC TTTCAGAAGC CGGCACTGGC CGTGGCGTGT TTTCTCAATG 360  
 GCGTGGCCAG CCGTGTGATG CTCTGCGGCT ACGCGACCTT CATTGCCAGC GTTCCGCCCA 420  
 TGTACACAG CATTGTGGCC TGTGCTGGG TGTCCCTCAA TGCATGGTCT TGGTCCACAG 480  
 TTTTCCACA CAGCGGACCT GACCTACAG AGAAATGGA CTACTCTGT GCTTCACATG 540  
 TCACTCTACA CTAATCTAC CTGTGCTGGG TCAGGACCGT GGGGCTCAAG CACCGACGCT 600  
 TGTGTCAGTG CTTCCGGGCT CTCTGCTGCG TCATGCTGAC CGTGCACGTC TCTCATCTGA 660  
 GCCTCATCGC CTCTGACTAT GGCTCAACCT TGGTGGCCAA COTGGCTATT GGCTGTGTCA 720  
 AGCTGGTGTG GTGGCTGGCG TGTGCTGTGT GBAACAGCG GCGGCTGCT CAGTGGGCA 780  
 APTGCTGTGT GTGTGCTGTG TGTGCTGAG GGTCTGCTCT CCGTAGGCTG CTTGACATTC 840  
 CACCGCTCT CTGGGCTCTG GATGCCATG CATTGCTGAG CATACGAC ATCTCGTGT 900  
 ACGTCTCTTT TCTGACGTTT CTGGAAGATG CACTCGCTGT CACTCTGAGG GAATCAGAG 960

ACAAATGCAA GCTGGACITGA AGACCTTGGG GCGAGTCTGC CCGAGTGGGG ATCTGCCCC 1020  
 GCGCTCTGCT GCTCTCCCTT TCCCTCAAC CTTTGAGATG ATTTCTCTT TCAACATCT 1080  
 TGAATCTGGA CATGAAGGAT GTGGGCCCCA AATCATGTGG CCAAGCCACC CCTTGTGGC 1140  
 CTTCAACAGC CTTGGAGTCT GTTCTAGGGA AGGCTCCCA GCATCTGGGA CTCGAGAGTG 1200  
 GCGAGCCCTT CTACCTCTCT GAGCTGAAGT GGGGTGGAAC TGAGTGTGCT CTAGCTCTTA 1260  
 CGGGAGAGAC AGCTGCGGCT TGTCTCCCA TGAGCTCTCT CCGACATCC CCAAGTCCCT 1320  
 GGGTGGGTC TGAAGCCCTC TGCTACTCTG GGAAGCCAGG GACCACAGGC CTTAGGGATA 1380  
 CAGGGGGTCC CTTCTTGTTA CCACCCCCCA CCTCTCCCA GGACACCACT AGGTGTGGCT 1440  
 GGATGCTGTG TCTTTTGCCA GCGAAGGTTC ACGGCGATTC TCCCCATGGG ATCTTGAGGG 1500  
 ACCAAGCTGC TGGGATGCG AAGGAGTTTC ACCCTGACCA TTGCTCTAGC CAGGTTCOCA 1560  
 GGAGGGCTCA CCAATCTGCT TTGAGGCGCC AGGCTGCGCC GCGAGCCAGG GCAAGATGTC 1620  
 TTGTCTGCTG TCTGTGTGAG AGCCTCCACC CGTGTGTCCG GAGTGTGGGC CAGGCTGAGT 1680  
 GCATAGGTGA CAGGGGCCGT AGCATGGGCC TGGGTGTGTG TGAGCTCAGG CTAAGTGGC 1740  
 CAGTGTGGAG ACGGGTGTGT TCGGGGAAGA GGTGTGGCTT CAAAGTGTGT GTGTGCAAGG 1800  
 GTGTGGGTGTG TTAGCCTGGG TTAGGGGAAC GTGTGTGGCG GTCTCTGTGG GCA TGTGGA 1860  
 TGAATGACTG CCGGTGAATG TGTCCACAGT TGAGAGGTGTG GAGCAGGATG AGGAATCTT 1920  
 GTACCATCA ATAAATGACT GTGAGGCGCC AGCTCTGGCC AAGGCCACAC CTGGGGGAGC 1980  
 AGCCAGGAGC TCTCCATGGC CAGGCTGGCT GTGTGCAAT TCTCTGTCTG GTGCCCTTT 2040  
 GCGCGCTCC TGCAAACTTC ACAGGGTCCC CACACAACAG TGCCCTCCAG AAGCAGGCC 2100  
 TCGAGGCGAG AGGAAGGAAA ATGGGGATGG CTGGGGCTCT CTCATCTCT CTTTCTCT 2160  
 TCTGTGGA TGCTGTGGCT TCCCTCCAAA AACCTCCATT CCGCTGTGC CAGGCCCTTT 2220  
 GCGATGCTCT GATTTTGGG AGGAGGAAGG GCGATTTTGA GGGAGAAGGG GAGAAATCTT 2280  
 ATGGCTGGGT CTGGTTTCTT CCGTCCCAAG AGGCTCTTAC TGTTCCAGG TGGGCCAGC 2340  
 GCAGGCGAGG GCGACACTAT GCTTGGCCCT TGTAAAGGT GACCCCTGCG ATTTACAGC 2400  
 AGCCCTGGCA TGTTCTGGC CCACAGGAAT AGAATGGAGG GAGCTCAGA AACTTCCAT 2460  
 CCGAAAGGCA GTTCTCGTGG TTGAAGCAGA CTGGATTTT GTCTGCGCC TGACCCCTGT 2520  
 TCCCTCTTG AGCTGAGGCT CTGAGGAG GATCTACA TGAGGCTGCT GTTTTGGT 2580  
 GCTAGCTTC TTTTGATCT GAAAACTTTT AAGTGTGGAG GGTGGCAAGG GATGTGCTTA 2640  
 ATAAATCAAT TCCAGGCTC AAAAAAAAAA AAAAAAAAAA AA

Seq ID NO: 50 Protein sequence:  
 Protein Accession #: NP\_219487.1

1 11 21 31 41 51  
 MKDVGPESCG QPTPCWPSA LESVLGKASQ HILBSQQL YLLELWGGT ECLSLSTGRT 60  
 AACPFISLS PPSAAWLOF EALCLPKRG TIGLIDTGGP LLLPPYLLQ DTRRWCMWLV 120  
 LWPVKVHDS FVCHLKDQA GCGKPEVDH CPKVPVRPH HPPFQDQSS KFRALRCC 180  
 LVESLPCVG SVQAGCEGD RAVSMGLGVC ELRPKCAVWR RVLGKRKGF KVCVRGWVC

Seq ID NO: 51 DNA sequence XM\_059089.1  
 Nucleic Acid Accession #:  
 Coding sequence: 178-518 (underlined sequences correspond to start and stop endons)

1 11 21 31 41 51  
 GATGTACACT CTGAAGTGA CACATCTG TTGGCAGGAC ATGACACCTT GGCACGAAGC 60  
 ATCTCTGGA TCTTTTACIT CTTGGCTCTG AACCTGAGC ATCAAGAGAG ATGCGGGAG 120  
 GAGTTCAGGG GCATCTGGG GGAATGGGCT TCTACTACTT GGGACAGCT GGGTGAAGTG 180  
 TGTATACCA CAATGTGCAT CAAGGAGAGC TGCCGATTGA TTCTGCAGT CCGCTCCATT 240  
 TCCAGAGATC TCAAGAGCC ACTTACCTTC CCAGATGGAT GCACATTGCC TGCAGGGATC 300  
 ACCGTGTGTC TTAATATTGG GGGTCTTAC CACAACCTGC CTGTCTGGAA AAACCCAAAG 360  
 GTCTTGACC CTTGAGGCT CTCAGGAGG AATCTGATC AGAGACACC CTAATGCTAC 420  
 TTACCATCT CAGCTGGATC AAGGAACCTG ATTTGGCCAGG AGTTGGCAT GATTGAGTGA 480  
 AAGGTAACCA TGTGCTGAT TCTGCTCAC TTCAAGATGA CTCAGAGCC CACCAAGGCT 540  
 CTACTCTTC CCAACCAATT TATCTCAA GCGAAGAATG GGATGTATT GCACCTGAAG 600  
 AAATCTCTGC AATGTLAGAT CTCAGGATAC AATGATTAAG CTGACTTTGT TTTTCGAAGT 660  
 TAAATTTACA GATATGATC CAAGCAGATA GAAGGAGATC AATGATATGG GGGAGATTG 720  
 GAGGTTGGTG GGAATAGGCT CTCTGAGGA AGATCAAAA TCAATTTCTG GTACACAGTG 780  
 GTGACGTAG ATCTGTCTCT ATATACTTT GGGAGATTIT CAGATCTITT CTGTAAATAT 840  
 TTCACTACTA TTAATGCTGT ATACACCAAT AGACTTTTAT ATATTTCTG TTGTTTAAA 900  
 AATAGTTTTC AGAATTATGC AAGTAATAAG TGCAATGTATG CTCAGTGTCA AAAATTCOCA 960  
 ACATAGAGAA ATCATGTAGA ATAAAAATT TAAATCTCAC TTCAGTTAGC CGACATTCCA 1020  
 TCCCTGAGC AATCTACT CTCTTCTCAA AAGCAGATA ATTGGGTG GATCTTTCA 1080  
 GACTTTTTC TATACATTT ATATGTAGAA ATGTAGCAAT GTATTGTAT AGATGTGATC 1140  
 ATCTATAT TGTAATGAT TTTTTCAC TAATAAATA TCACTTAT CTT

Seq ID NO: 52 Protein sequence:  
 Protein Accession #: XP\_059093.1

1 11 21 31 41 51  
 MSYTTMCIRK TCRIDPVPVS ISRDLKPLT FPDGCTLPAG ITVVLSTWGL HHNPVAVWKNP 60  
 KVDPLNRC RNSDHPHYA YLPFASGRSN CIGQFAMIE LKVTIALILL HPRVTPDPFR 120  
 PLTPNPHFL KPKNQMYLHL KKLSEC

Seq ID NO: 53 DNA sequence  
Nucleic Acid Accession #: NM\_030916  
Coding sequence: 1-1533 (underlined sequences correspond to start and stop codons)

5  
1 11 21 31 41 51  
ATGCCCTGT CCTGGGAGC CGAGATGTGG GGGGCTUAGG CCTGGCTGCT GCTGCTGCTA 60  
CTGTGCGCAT CATTACAGC CCGTGGCCCC GCGGTGTAGC TGGAGACCTC AGAAGTGGTA 120  
ACGTGTGTGC TGGGCGCAGA CGCAAAACCT CCTCGCTTCT ACCGAGGGGA CTCGCGCGAG 180  
CAAGTGGGCG AAGTGGCATG GCTCTGGGTG GACCGGGGCG AAGGCGCCCA GGAATCAAGC 240  
CTACTGACT CCAATGAGG GTTCTATGTG AGCCCGGCTT AGAGAGGCGC CDTGAGAGAC 300  
CGCCGCCGCC CAGCAACACC CTCTGACGGC TCACTGTCTC TGGCAACGCG AGTGCAGGCG 360  
GATGAGGGCG AGTACGAGTG CCGGGTCAAG ACCTTCCCGC CCGGAGGCTT CCAGGCGCGG 420  
CTGCGGCTCC GAGTGTGTGT GCTTCCCTGT CCTCACTGTA ATCTGTGTC AGCACTAGAA 480  
GAGGGCCAGG GCTTGACCTT GGCAGCTCC TGCACAGCTC AGGGCAGCCG AGCCGCCAGC 540  
GTGACCTGGG ACAGGAGGT CAAAGGACAA AGTTCAGGCC GTTCTTCAA GCACTCCCGC 600  
TCTGCTGCCG TCACCTCAGA GTTCCACTTG GTGCCAGGCC GCAGCATGAA TGGGACGCCA 660  
CTGACTTGTG TGGTGTCCCA TCTGGGCTG CTCAGGAGCC AAAGGATCAC CCACATCTCT 720  
CACGCTGTCT TCTTGTCTGA GGCCTCTGTG AGGGGCGCTTG AAGAACAAAA TCTGTGGCAC 780  
ATTGGCAGAG AAGGAGCTAT GCTCAAGTGC CTGAGTGAAG GGCAGGCCCC TCCCTCATAC 840  
AACTGTGACG GGTCTGATGT GCTCTTGCCC AGTGGGTAAC GAGTGTGATG GGCATCTATG 900  
TGCTTTCCCG CAGTACGACC TGAAGCAGCG GGCATCTACG TCTGCTATGT CAGCAATGAG 960  
TCTCTCTCAA GGGATTCTCA GTTCACTGTG GATGTCTTGT ACCGCCAGGA AGACTCTGGG 1020  
AAGCAGGTGG ACCTAGTGTG AGCTTCGGTG GTGGTGGTGG GTGTGATCCG CGACTCTTGT 1080  
TCTGCTCTTC TGGTGGTGGT GGTGGTGTCT ATGTCCGAGT ACCATCGGCG CAAGGCCAG 1140  
CAGATGACCC AGAATATGTA GAGGAGAGCTC ACCCTGACCA GGGAGAACTC CATCCGAGAG 1200  
CTGATGCCCT ATCAAGAGAG CCGTCAAGAC AGAGTCTGAG GCTGAGAGCG 1260  
GAGGCGACCC CTGATAGTCT CAAGGACAA AGTAGTGTCT CTGTGATGAG TGAAGAGGCC 1320  
GAGGGCGCGA GTTACTCAC GCTGACCAAG GTGAGGGAGA TAGAAACACA GACTGAAGCT 1380  
CTGTCTCCAG GCTCTGGGCG GCGCGAGGAG GAGGAAGATC AGATGAAGG CATCAACAGC 1440  
GCCATGAACC ATTTTGTCTA GGAAGATGGG ACCCTACGGG CCAAGGCCAC GGGCAATGGC 1500  
ATCTACATCA ATGGGCGGGG AGAAGCTGTCTTGA

Seq ID NO: 54 Protein sequence:  
Protein Accession #: NP\_112178.1

1 11 21 31 41 51  
MPISLGAEMW GPFAWLILL LLSFTGRCP AGELEISDVV TVVLQDAKL PCFYRGDSGE 60  
QVGQVAWARV DAGEGAQLA LLHSKYGLHV SPAYEGRVEQ PFPNRPLDG SVLLKNAVQA 120  
DHGEYECRVS TFPAGSPQAR LRLRLVPLP PSLNPGPALE EGQLTLAAS CTAEGSPAPS 180  
VTWDTVEVGT TSSRSFKHSR SAAVTSEFHL VPSRSMNGQP LTCVVSHPGL LQDKRTHLL 240  
HVSFLASVS RGLDQNLVH IGRDAMLEK LSRSGPPSY NWRLDPLP SGVIRDQDL 300  
GFPLLTTHS GIYVCHVNS PSRSRSQVTV DVLDQEDSG QKVLDVLSAV VVVGVIALL 360  
FCLLVVVVVL MRSYHRRKQA QMTQKYBEE LLTRENSIKR LHSHHTDPS QFESVGLRA 420  
GHPHPSLKDIN SCSVMSSEP EGRSYTLTT VREIETQTEL LSPGSGRAEB EEDQDEGIQ 480  
AMNHVPQENG TLRAKPTGNG IYNGRHLV

Seq ID NO: 55 DNA sequence  
Nucleic Acid Accession #: AF007170.1  
Coding sequence: 73-1725 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
AAGGAGGGCG CCTCGGGGAA AAGGACCGC AGGACTCTG AGAGCAGCCT CCATGAGGCC 60  
CTGGACCACT GCAATGACGC CTTGGACCTC TTCTTACCA ACCAGTTCTC AGAAGCACTC 120  
AGCTACTCTA AGCCAGACA CAAGGAAAGC ATGTACCACCT CACTGACATA TGGCACCATC 180  
CTGAGATGCG AGCCCAATGAT GACCTTTGAC CTTCAAGGACA TCTGCTTGC CGGCAACAAT 240  
ATAAGAGAGG CACAGATGCT GTTCAAGAGG CACCGGAGGA AGTCTTCTGT AACAGATCC 300  
TTCAAGAGCC TTGTGACGCG GCGCAAGCTG GCGCAATCA CTGAGAGAGA AATCAAGTCA 360  
GAGGTCTGCT ATGCAGATGT CCTGTGCAAG CGAGCAGGCC TGACCTTCTT GCAGAGCAGG 420  
AACATGTGTA GCTTACTCAA AGGCGGCATC AAAGTTGAAA ACAGCTACCA GACCTACAAG 480  
GAGCGTGACA GCTTGTITCA GCTCTCACA TACTGCAAGG GTGAGAACCA CCAGCACTTT 540  
GAGGAGGAGG TGAAGCTTGG TGTAGGGGCG TTCAAGCTTA CACTGTCCAT GCTTCTACT 600  
AGGATCTGTA GCGTGTGTA GTTTGTGGGG TTTCACGAAA ACAARAATA TGGGTGCTG 660  
CAGCTGAGAG AGGGAGGCTC AGGBCACAGC TTCCGCTCTG TGTCTTGTGT CATGTCTCTG 720  
CTGTGCTACC ACACCTTCTT CACCTTCTGT CTGGTACTG GGAACGTCAA CATGAGGAG 780  
GCCGAGAAGC TCTTGAAGCC CTACCTGAAC CGTGAACCTA AGGTTGCCAT TCTCTGTTTC 840  
TTTGCAGGGA GGAATGAGAT CATTAAGGCG AACATTGATG CAGCAATCCG CGGTGTGGAG 900  
GAGTGTCTGT AGGCGAGCA GCACTGAGAG CAGTTTCACC ACATGTCTTA CTGGAGCTG 960  
ATTGTGTGCT TCACCTACAA GGGCAGATGG AAGATTGCTT ACTTCTACCG CGACTCTCTC 1020  
AGCAAGGAGA ACTGTGCTG CTAGGCCACC TACATTACCA TGAAGGCCCG CTACCTCAGC 1080  
ATGTTTGGGA AGGAGGAGCA CAGCCGCTTC GGGGAGCAGC AAGTGGAAAT ATTTGAGAGT 1140  
GTGCGAGGGA TGAAGCTCAA GTTACTTGGG AAATCTCTAC CCACAGAGAA GTTTCGATC 1200  
CGGAAGTCCC GGGCTACTT CTGCTGAAC CCTATCTGCG TGCCATAGCC TCGCTGTGAA 1260  
ATGATGTACA TCTGTGAGC CTACGCGGTG ATTGGGAAGC AGCCGAACCT CACGAGTGGG 1320  
ATACCTGAGA TTATCACTAA GGCTGAAGAG ATGTGTGAGA AAGGCCAGGA GAAGGATGAC 1380

TCAGTGGATG ACGAGTGCCT GGTGAAATTTG TTGAAAGGCC TGTGTCTGAA ATACCTGGGC 1440  
 CQTGTCAGG AGGCCGAGGA GAATTTTATG AGCATCTCTG CCAATGAAAA GAAGATTAAA 1500  
 TATGACCACT ACTTGATCC AACGCCCTG CTGAGCTGCG CCTGCTGCT TATGGAGCAA 1560  
 GACAGAAAGC AAGAGGCCAT CAACATTTTG GAATCTGCCA AGCAAACTA CAAGAATTAC 1620  
 TCCATGGAGT CAAGGAACACA CTTTCGAATC CAGGCAGCCA CACTCAAGC CAAGCTCTCC 1680  
 CTAGAGAACA GCACAGATC CATGGTCTCA TCAGTGTCTT TGTAGCTTTG TGCAGCAGTT 1740  
 CCGGCTGGA AGACAGAGAC AGCTGGACAG AGCTCTCTGA ACATTCTCAA AATACCCCTT 1800  
 CCCCCTGCC TGCCTGACCT TTGGGTGTCA CCGGCACCTC AGTTGGATGG CACAAATAG 1860  
 TGTATCCGTG CAGAAGCCGA GCTGGCATT TCAACAGTGT AGCCAAGGCC CTTTGGCAAC 1920  
 GGCAGAGCAG GTGGAGCCCT CTGCTTGCCC TATCACACAT AGCGGTACTT GCTTTTCACT 1980  
 GTGATGTTTA AGAGAAATGA TGAACAGTTT ACATTTTCTT TAGAAATACA TTAGTGGGAT 2040  
 CACAGTGTGC TTAAAAACC AACAAZATC AACCCGCTT AAGTCTGT AGTCAATAT 2100  
 TATGATCTGG AGGTAATCTT CTTTATATGA TGTAGCCAAA GGCACAAATG CTTTCAAAAT 2160  
 TCAGCAAGTT CTCAGCTTGT GTGACGGGAG GTCTCTCAGA GGACCTGAGG AATGCTGGG 2220  
 AGAGGCTAAG CCTCAGCCTT CAATGCTCTT GGGGTGTGGC ATGAGGATGT ACACAGACAC 2280  
 CCACATCCTT ACTACTCACA CTCATTCTCA CTCCTTTTGT AAATTTCCAA TTTAAAAATC 2340  
 AAGCAGCTCT TTTATGTAG ATAAATCTGT AGCTCTTCTG TAGAATAATC AATGCTTACC 2400  
 AGTAGAAAT GCGAGGGCTT GATGGAAGAG CAGTGTATGC CTTTCTATGC CAAGAAGCAG 2460  
 AAATTTGGGG GGCAGGAGGA GGTTCCTAGA ATCCAGTCTG TATCTTGTCT GTATGCCAAA 2520  
 CTGAACACAC TGGGATAAT TATGAAACA TAAAAATCTT CTGATCTCA CTCGAAAGTA 2580  
 CATTGCTTA CTCAGCAGAT TTTTGTTAAA ACTGTTATTC TTGAAAAAAA AAAAAAATAA 2640  
 AA

Seq ID NO: 56 Protein sequence  
 Protein Accession #: AAC39582.1

1 11 21 31 41 51  
 | | | | |  
 MTALDIFLNT QFSEALSVLK PRTKESMYHS LTYATILEMQ AMMTDFDQDI LLAGNNMKEA 60  
 OMLCRHRRK SSVTDSFSL VNRPTLGGT EEEHAEVCY AECLQRAAL TFLQDENMYS 120  
 PIKGGIKVRN SYQTYSKELDS LVQSSQYCKG ENHPHFBOG KLVGAPNLT LSMPLRILR 180  
 LLRFVGFSON KDYGLQLQLE GASGHSFRSV LCVMLLLCYH TPLTFVLGT NNNIEBAEK 240  
 LKPYLNRYPK GAFLFPFAGR IEVKGNDIA ABRPEECCE AQHWKQFHH MCYVSLMWPC 300  
 TYKGQWMSY PFDILSKEN CWSLATTYTM KAAYLSMFCR EDRKPPDDH VELRAVPL 360  
 KLRKAGSLP TEKFAURSKS RYFSSNPISL PVPALBMMYI WNGYAVIGVK PKLTGILEI 420  
 ITKAEEMLEK GPENIYSVDD ECLVLLKGL CLKYLGRVQE ABENFRSIA NEKKIKYDHY 480  
 LIPNALLELA LLLMBQDRNE EAKLLESAS QNYKNYSMES RTHFRQAAT LQAKSSLENS 540  
 SKSMVSSVSL

Seq ID NO: 57 DNA sequence  
 Nucleic Acid Accession #: NM\_006670.1  
 Coding sequence: 1-927 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 | | | | |  
 ATGCTCTGGG GGTGCTCCCG GGGCCCCGCG GCGCGGGAAG GCGCTCTCG GCTGGGCGGA 60  
 CTAGCGCTGG TACTCTGGG CTGGGTCTCC TGTCTTCTC CCACTCTCTC GGCACTCTCC 120  
 TTCTCTCTCT CGGCGCGGTT CTGCGCTTCC GCGGTGTCGG CCGAGCCGCC GCTCGCGGAC 180  
 CAGTGCCTCC CGCTTGTGGA GTGCTCCGAG GCAGCGGCGCA CAGTCAAGTG CTTTAAACCG 240  
 AATCTGACG AGTCTCCAC GGCATCTGCC GCTCAGTGGC GCAACTCTTT CCGTACCGCG 300  
 AACCGCTGGC CCGACCAACA CTCCTTTAC CTGCGCGCGG ATGTGCTGGC CCAACTGCCC 360  
 AGCCTCAGGC ACCGTGACCTT AAGTAATAAT TCGTGTGTGA GCTGACCTA CGTGTCTTCC 420  
 CGCAACCTGA CACATCTAGA AAGCCTCCAC CTGGAGGACA ATGCCCTCAA GGTCTCTCAC 480  
 AATGGCAACC TGGCTGAGTT GCAAGTGCTA CCCCACATTA GGGTTTCTCT GGACACAAAT 540  
 CXTGTGTGTT GCGATGAGTA CAGTGACGAG ATGCTTAACCT GGCCTACAGA AACAGAGTA 600  
 GTGACGAGCA AAGACCGGCT CAGCTGTGCA TATCCGGAAA AAATGAGGAA TCGGGTCTCC 660  
 TTGGAACCTA ACAAGTGTCTGA CTTGGACTGT GACCCGATTC TTCCCGCATC CCTGCAAAAC 720  
 TCTTATGTC TCTTGGGATAT TGTITTAGGC CTGATAGGCG CTATTTCTT CCTGGTTTGG 780  
 TTTTGAAGC GCAAGGGGAT AAAAAAGTGG ATGCATAACA TCAGAGATGC CTGACGGGAT 840  
 CACATGGAAG GGTATCATTA CAGATATGAA ATCAATGGCG ACCCCAGATT AACAAAGCTC 900  
 AGTTCTAACT CGGATGCTCT CGAGTGA

Seq ID NO: 58 Protein sequence  
 Protein Accession #: NP\_006661.1

1 11 21 31 41 51  
 | | | | |  
 MPGGCSRGPA AGDGRRLRAR LALVLLGWVS SSSPTSSASS FSSSAPLAS AVSAQPLPD 60  
 QQCALCBCE AARTYQVKNR NLTEVPTDLP AYVRNLFLE NQLASNUFLY IPRDLVQLP 120  
 SLRHLDLNN SLVSLTYSP RNLTELESL LEPNALVKLY NGLTARQL PHRFVLDN 180  
 PWFVCGEDMAD MYTLWKETEV VQGKDRITCA YPKMRNRVL LELNSADLDC DPILPSSLQT 240  
 SYVFLGVLA LIGAFLLVL YLNRKGKZKW MHNRIDACRD IMEGSYHYRE INADPRITNL 300  
 SSNSDVLE

Seq ID NO: 59 DNA sequence  
 Nucleic Acid Accession #: NM\_024022



Coding sequence: 1-1362 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 5 ATGGGGGAAA ATGATCCGCG TCGTGTGTGA GCGCCCTTCT CATTCGGATC GCTTTTGGC 60  
 CTGATGATTT TGAATAAAG TCGTGTGTGA CCAGATGCG AGTCGTGTC TGACAGATG 120  
 CTGTCACTCG TGCCATGAAA GTTTTTCAT ATCATGTCAT TTGGGATCAT TGCAATGATA 180  
 TTAGCACTCG CCATGTGCTT GGGCATCCAC TTCGACTGCT CAGGGAAGTA CAGATGTGCG 240  
 TCATCTTTTA AGTGTATGCA GCTGATAGCT GATGTGACG GAGTCTCGGA TTGCAAAAGAC 300  
 GGGGAGGACG AGTACCGCTC TGTCCGGGTG GGTGTGTGTA AGTCCGTGCT CAGAGTGTTC 360  
 ACAGTCTGCT CTGTGAGAC CATGTGCTCC GATGACTGGA AGGGTACTTA CCGAAATGTT 420  
 CGCTGTGCCC AACTGTGTTT CCCAAGCTAT GTGAGTTCAC ATAACCTCAG AGTGAAGCTCG 480  
 GTGGAGGGGC AGTTCGCGGA GGAAGTTGTG TCCATGTCAT ACCTCTTGCC AGATGACAAG 540  
 GTGACTGCAT TACAACCTAC AGTATATGTG AGGAGGGATG GTGCCTCGG CCAAGTGTGT 600  
 ACCTTGCACT GCACAGCTTG TGATCATAGA AGGGGCTACA GCTCAGCAT CTGTGGTGTGA 660  
 ACATGTCCT TCTCTTGCGA GTGGCCCTGG CAGCGACGCG TTCACTTCCA GGGTCAACAC 720  
 CTGTGCGGGG GCTCTGTGAT CAGCGCCCTG TGGATCATCA CTGCTGCACA CTGTGTTTAT 780  
 GACTTGTACC TCCCAAGTC ATGGAACATC CAGGTGGGTC TAGTTTCCCT GTTGGACAAT 840  
 CCAGGCCCAT CCCACTGGT GGAGAAGATT GTCTACCACA GCAAGTACAA GCCAAAGAGG 900  
 CTGGGCAATG ACATCGCCCT TATGAAGCTG CCGCGGCCAC TCACGTTCAA TGAATATGATC 960  
 CAGGCTGCTT GCTCTGCAAA CTCTGAGAGT GACTTCCCGC ATGGAATAAT GTCTGGAAG 1020  
 TACGATATGG GGGGACAGAA GATGTGAGGT GAGCGTCCCT CTGTCTCGAA CCACGGGGCG 1080  
 GTCCCTTTGA TTTCACAACA GATCTGAAC CACAGGAAAG TGTAAGGTGG CATCATCTCC 1140  
 CCGTCCATCG TCTGGCGGGG CTACCTGACG GTTGGCGGTG ACAGCTGCCA GGGGAGCAAC 1200  
 GGGGGGCCCC TGTGTGTGTA AGAGAGGAGG CTGTGGAAGT TAGTGGGAGC GACCACGTTT 1260  
 GGCATCGGCT GCGCAGAGGT GAACAAGGCT GGGGTGTACA CCGGTGTGAC CTCTTCTCTG 1320  
 GACTGGATCC ACAGACGATG GGAAGAGAGAC CTA AAAAAGCTGA

Seq ID NO: 60 Protein sequence  
Protein Accession #: NP\_076927

1 11 21 31 41 51  
 35 MGENDPPEV EAPFSSRLFG LDDLKISPIA PDADAVAAQI LSLPLKFFP IVIGHALI 60  
 LALALGLGH FDCSGKYRCR SSFKCIELA RCGVSDCKD GEDEYRCVRV GGQNAVLQVF 120  
 TAASWKTMC SDDWKGHYANV AACAQLGPSY VSSDNLNRSV LBGQFREEFV SIDHLPLDPK 180  
 VTALHHSYIV NGCAAGTCRRA RGTSTAGGIR ROTSSSRIVG NMBSLQZPWF GSCZQFYOTI 240  
 LGGSSVTPLE WIYAAAGYPI DYLRKSWTI QVGLVSLLEN PAPSILVEK VYHBSYKPKR 300  
 LQNDIALMKL AGLPTFNEMI QPVCLPNSSE NFPDGGKVCWT SGWGATEDGG DASPLNIFIA 360  
 VPLSNKICN HRDVTGGIIS PSMCLCAGYL GTVDSCQDGS GGPLVCQERR LWKLVGATSF 420  
 GIGCAEVNKP GVTYTRVTSF DWIHEQMERD LKT

Seq ID NO: 61 DNA sequence  
Nucleotide Accession #: NM\_006475  
Coding sequence: 28-2538 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 50 AACAGAACTG CAACGGAGAG ACTCAAGATG ATTCCTCTTT TACCACATGT TTCTCTACTA 60  
 TTGCTGCTTA TGTGTAACCC TATAAAGGCC AACCAACTATT ATGACAAGAT CTTGCTCAT 120  
 AGTCGTATCA GGGGTGCGGA CCAAGGCCCA AATGTCGTGT CCCTTCAACA GATTTTGGCG 180  
 ACCAAAAGAA AATACTTCAG CACTTGTAAAG AACTGGTATA AAAAGTCCAT CTGTGAGACG 240  
 AAAACGACTG TTTTATATGA ATGTGCGCTT GTTTATATGA GAATGGAAGG AATGAAAGCG 300  
 TGCCGAGTGA TTTGAGCTTA TGACCAATGTT TATGACACTG TGCGGACTGT GGGAGCACC 360  
 55 ACAACGACAG GCTATTCTGA CCGCTCAAAA CTGAGGGAGG AGATCGAGGG AAAGGAGATCC 420  
 TCTCACTACT TTGACCCGAG TAATGAGGCT TGGGACAACCT TGGATTCTGA TATCCGTAGA 480  
 GGTTTGGAGA GCAACGTGAA TGTGTAATTA CTGATGCTT TACATAGTCA CATGATTAAT 540  
 AAGAGATATG TGACAAGAGA CTTAAAAAAT GGCATGATTA TTCTCTCAAT GTATAACACT 600  
 TTGGGCTCT TCAATTAACA TTACTCAAT GGGTGTGTCAT CTGTAATTG TCTGCACTAC 660  
 60 ATCCATGGGA ACCAATATTC ACAAAATGCT GTGTGCTCATG TCATTGACCG TTGTGCTATA 720  
 CAAATTTGTA CTTCAATCTA AGACTTCATT GAAGCAGAAAG ATGACCTTTC ATCTTTTGA 780  
 GAGCTGCCCA TCACATCGGA CATATTGGAG GCGCTTGGAA GAGACGGTCA CTTACACTC 840  
 TTTGCTCCCA CCAATGAGCG TTTGAGAAA CTTCCAGGAG GTGTCTTGA AAGTGTTCATG 900  
 GGAAGCAAGT GAGTCTTCCA AGCTCTTATG AAGTACACAA TCTTAAATCT TCTCAGTBT 960  
 65 TGTGAGTCTA TTAGGGGAGG AGCAGTCTT GAGAGCTGG AAGGAAATAC AATTGAGATA 1020  
 GGATGTGACG GTGACAGTAT AACAGTAAAT GGAATCAAAA TGGTGACAAA AAAGGATATT 1080  
 GTGACAAATA ATGTGTGTAT CCAATTGATT GATCAGGCTC TAATTCTGA TCTGCGAAA 1140  
 CAAATGTTAT AGCTGGCTGG AAAACAGCAA ACCACTTCA CGGATCTTGT GGGCCAAATA 1200  
 GGCTTGGCAT CTGCTCTGAG GCGAGATGGA GAATACACTT TGTGCGCAC TTGTAATTAAT 1260  
 70 GCATTTCTGT ATGATGCTC CAGCATGGTT GAGCGGCTCC TTAATAATAT TTGTGAGATA 1320  
 CACATATTGA AAGTAAAAAT TGGCTTAAAT GAGCTTTACA ACGGGGCAAT ACTGGAAGAAC 1380  
 ATGCGAGGCA AACAGCTCAG AGTCTTGTA TATGTGACAG CTGTCTGCAT TGAATAATCA 1440  
 TACATGGAGA AAGGGAGTAA GCAAGGGAGA AACGGTGCGA CAGATATCTT CCGCGAGATC 1500  
 ATCAAGCCAG CAGAGAAATC CCTCATGAAA AAGTGAACAC AAGATAAGCT CTTTGGACCC 1560  
 75 TTCTCTGCT TACTGTGCTG TCGAGACTGA AAGAGCTCC TGACACACC TGGAGACTGG 1620  
 ACATTAATTT TGGCAACAA TATGCTCTTT AAGGAGAGGA CTAGTGAAGA AAAAGAAAT 1680  
 CTGATACGGG ACAAAATATG TCTTCAAAAT CATACTCTTT ATCACTGAC AACACAGGAT 1740

TCTATTGGAA AAGGATTGA ACGTGGTGT ACTAACATTT TAAGAACCC ACAAGGAAGC 1800  
 AAATCTCTTC TGAAAGAAAT AATGTATACA CTTCCTGGTA ATGAATTTGAA ATCAAAAGAA 1805  
 TCTGACATCA TGACAAACAA TGGTGAATTT CATGTTGTAG ATAAACTCTT CTATCCAGCA 1920  
 GACACACCTG TTGGAATGA TCAACTGCTG GAAATACTTA ATAAATTAAT CAAATACACT 1980  
 CAAATTAAGT TTGTTGGTGT TAGCACTCTT AAAGAAATCC CGGTGACTGT CTATACAACT 2040  
 AAAATTATAA CAAAGTGTG GGAACCAAAA ATTAAAGTGA TGGAAAGCAG TCTTCAAGCT 2100  
 ATTATCAAAA CTGAAGGACC CACATCAAAA AAGATCAAAA TTGAAGGTGA ACCTGAATTC 2160  
 AGACTTGATA AAGAAGGTGA AACAAATACT GAAGTGATCC ATGAGAGGCC AATTATATAA 2220  
 AAATACACCA AAATCATTTA TGGAGTGCTT GTGGAAATAA CTGAAAGAGA GACACGAGAA 2280  
 GAACGAATCA TTACAGGCTC TGAATAAAA TACACTAGGA TTTTACTGTG AGGTGGAGAA 2340  
 ACAGAGAGAA CTCTGAAGAA ATTGTATACA GAAGAGGTGA CCAAGGTTCAC CAAATTCATT 2400  
 GAAGTGTGTG TGTGAACATTT ATTGAAGAT GAAGAATAA AAAAGACTGT TCAGGAGAGC 2460  
 ACACCGGTGA GSAAGTTGCA AGCCACAAA AAGTTCAGAG GTTCTAGAAAG ACCTAATAAG 2520  
 GAAGTGCTGT TCAAGTGA AAATCCAAAA CAGAAAAAAA TGTTTATACA ACCTAATAGT 2580  
 AATAACCTGA CCTTAGAAAA TTGTGAGAGC CAAAGTTGACT TCAGGAAGCT AATCATCAGC 2640  
 ACAGAGAGAG AATCATCAAA TAATTCTGAA CACAAATTTA ATATTTTTT TCTGAATGA 2700  
 GAACATGAG GGAATTTGT GAGTTAGCTT CCGTGTGTAA AGGAATTGAA GAAATATATA 2760  
 CACCTTACAC CCTTTTTCAT CTGACATTA AAGGTTTGTG CTAAGTTTGG AATCCATTA 2820  
 AGAAAAATC TTGTACACAG ATTCAATCA ATTCAAATCG AAGAGTTGTG AACGTTCAT 2880  
 CATTGAAAA GACCCAGCTT TGTATGATG TTATGGATAC ATAAATGCA CGCAAGCCAT 2940  
 TATCTCYCA TTGGGAAGCTA AGTTATAAAA ATAGGTGCTT GGTGTACAAA ACTTTTTA 3000  
 TCAGAAAGCT TTGCACATTT CTATATGAGT GGGTTACTGT GTAAATTATG TTAATTTTA 3060  
 CAACTAAAT TTGTACTCCA GAATGTTGT CATATGCTCT TCGCAAGTGA TGAATTTA 3120  
 TCTCAAGGT TTCAATAAAA CCATTTTTCAT GATATAAAGA GAATTAATCT AAATGAGTA 3180  
 ATTCAGAAAA ACTCAAGATT TAAGTTAAAA AGTGGTTTGG ACTTGGGA

Seq ID NO: 62 Protein sequence:

Protein Accession #: NP\_006466

1 11 21 31 41 51  
 MIPFLPMTSL LLLLVNPN ANHYDKILA HSRIGRQDQ PNVCALQL GTKKYFSTC 60  
 KNWYKSKSG QTTVLVECC PGYMRMBGMK GCPA VLPIDH VYGLTGLVGA TITQRYSDAS 120  
 KLREREGSG EFTYFAPNE A WNEI DSDIR KGLSESVVE LKALSHMI KRLMLTDEK 180  
 NGMIUPTSMN NLGLFNHPY NGVVTYNCA IHGNQIATN GVYHIDRVL TQGTSGIQF 240  
 IEAEDDLSSF RAAATSDIL EALRGDRGHT LFPATNEAFB KPLRGVLEF MDGKVASEAL 300  
 MKYHILNTLQ CSSEIMGGAV FETLEGNTE IGCQGDSTV NGKMKVNNKD IVTNNGVHL 360  
 ILQVLPIDSA KQVBLEAGK QTTIFDL VAQ LILASLRPD GEVTLAPYN NAFSDITLSM 420  
 VQRLKLLQ NHRQILK TIQGRKLVNF VYRTAVCIEN SKMRKSGQ 480  
 RNGAHLPBF IKPABKSLH BKLKQDKRHS TPLSLLEAD LKELLYQPD WLPVFNDA 540  
 PKGMTSEBKE ILIRDKNALQ NELYHLTFP VFPGKGFEPF VYNILKTQD SKILKEVND 600  
 TLLVNLKSK BSDIMTTNGV IHVVDKLLYP ADTPVGNQDL LEILNKLIY IQKQFVRGS 660  
 FKPEPVTYVT TKITKVPVP KIKVEGSLQ PIKTEGPTL TKVKEGEPF FLIKBQETI 720  
 TEVHGEPE KKYTKIDGV PVEITEKTR BERITGPEI KYRTSTGGG STEELKLL 780  
 QHEVTKVTFE IEGGDHLEF DEEKRLIQD DTPVYKLQAN KKVQSRRLR RBRGQ

Seq ID NO: 63 DNA sequence:

Nucleic Acid Accession #: NM\_020974

Coding sequence: 81-3083 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 GCGCTCCGGG CACACTCCCC GCGCGCGCGG CCGGCAACGC CCGCACTCCG CCGCCTCTGC 60  
 CCGCAACCGG TGAAGCATTC ATGCGGHTTC GCGAGCGGCA CGTCCCGGG CGGCGCTGGG 120  
 CCGTCTCTCT GTCTCTCTGT CTGCTGCGCC CACGCTGTCT GCTGCTCCCG GCGCTCCCG 180  
 CCGGTCGGGG CCGTGGCGGG GCGCGCGGAG CAGCTGTGCA TGAATGTGTG CAAAGGCTAG 240  
 ATGACTGCCA TGCAGCAGCC CTGTGTGAGA ACACACCCAC CTCCTACAAG TGCTCTCTGC 300  
 AGCTTGCTCA CCAAGGGGAA GCGAGGCAAT GTGAGGACAT CGATTGAATG GGAATGAGC 360  
 TCAATGGAGG CTGTGTCCAT GACTGTTTGA ATATTCCAAG CAATTATGTT TGCACATGTT 420  
 TTGATGGCTT CATGTGGCT CATGACGGTC ATAAATGTCT TGAATGGAG GAGTGGCTGG 480  
 AGAACATATG CCGCTCCGAC CATACCTGTT TCAACATGTT GCGAGCACTT GAGTGGCTG 540  
 CCAAGAGGCG GTTTTTCCTG AGTGACAAA CTAGCAACCTG CATTACCGCC TCGAAGAAGG 600  
 GCGCTGAGCTG CATGAATAAG GATCACGGCT GTAGTCACAT CTGCAAGGAG GCCCAAGGG 660  
 GCGAGCTGCG CTGTGAGTGC AGGCGTGGTT TTGAGCTGGC CAAGAACCAG AGAGACTGCA 720  
 TGTGTGAGCT TAACCATGGG AACGCTGGGT GCGACGACTC CTGTGACGAT ACAGCCGATG 780  
 GCGCAGATG CAGTGGATCAT CCAACATACA AGATGACAC AGATGGGAGG ACCTGCTCT 840  
 AGGAGAGGCA CACTCTGCTT GAGGTGACAG AGAGCAACAC CACATCACTG GTGTGATGGG 900  
 ATAAACGGGT GAAACGGCGG CTGCTCATGG AAACGTGTGC TGTCAACAAT GGAGGCTGTG 960  
 ACAGCACTG TAAGGATACT TGACAGAGTG TCAACTGCAG TTGTCTGTT GGAATCACT 1020  
 TCCAGTTGGA TGGGAAGACA TGTAAAGATA TTGATGATGT CCGAGACCGG AATGAGAGTT 1080  
 GTATCATTTT CTGCAAAAAC ATCTGCGTCA GTTTTATGTA CCGGTGAGCA AAGAGATTTA 1140  
 AATTATTAA AGATGAGAA GTCTTCCGAA ATGTGATAGA GTGCTTTTG GATAGGACT 1200  
 GTGACACAGC CTGACTCAAC CACCTCGGCA CATTGTCTGT TGCTTGAAC CGAGGATGCA 1260  
 CCTGTATGG CTTCACCCAC TGTGGAGACA CCAATGAGTG CAGCATCAAC AACGGAAGCT 1320  
 TGCAGCAGGT CTGTGTGAAC ACAGTGGGCA GCTATGAAT CCAATGCCAC CTTGGGTGCA 1380  
 AGCTCCACTG GAATAAAAA GACTGTGTGG AATGTAAGGG GCTCTGCGCC ACAAGTGTGT 1440  
 CACCCCGTGT GTTCCGACAC TGGGTGAGA GTGTGGAGG AGACGGGCTT TCTCTAGAT 1500

GTCACTCTGG CATTCACTTC TCTCAGATG TCACCACCAT CAGGACAAGT GTAACCTTTA 1560  
 AGCTAAATGA AGGCAAGTGT AGTTTGAAAA ATGCTAGCTT GTTTCOCGAG GGTCCTGCGAC 1620  
 CAGCACTAOC AGAGAAGCAC AGCTCAGTAA AAGAGAGCTT CGGTACGTA AACCTTACAT 1680  
 5 GTGAGCTCTGG CAAGCAGTGC CCAGGAGGCC CTGGGCGACC AAGCAACCCCT AAGGAAATGT 1740  
 TTATCACTGT TGAATTTTGG CTGGAAGCTA ACCAAAAGGA GGTCACAGCT TCTTTGTGAC 1800  
 TGAAGTCTAT GTTAAGAAGA ACCGAGAAAG GGCTCCTTAA AGCGATCCCG AGGCTGAGAA 1860  
 AGGCCGTCCA CAGGAGGACG TTTCACTCTC AGCTCTCAGG CATGAACCTC GAGCTGGCTA 1920  
 AAAAGCCTCC CAGAACATCT GAAAGCCAGG CAGAGTCTGT TGAAGTGGCG CAGGGTCACT 1980  
 10 CAGAAAACCA ATCTGTGCACT TGCAGGGCTG GGACCTATTAT TGAATGGAGCA CGAGAACCCCT 2040  
 GCAATTTTAT TCCAAATATGA ACCTTCCAAA ATGAGGAAGG ACAAATGACT TTGGAACCAT 2100  
 GCGCAAGACC AGGAATAATCT GGGCCCTGTA AGCTGGAAT ATGTGTGAAT 2160  
 GTGAGAGTCT GTTCAAGCTT GGTAATATT GTGCAGATGG GTTGAACCT TGCAGACCT 2220  
 GTGCCCTGGG CAGCTTCCAG CCTGAAGCTG GTGGAACCTC TCGTCTCCCC GTGGGAGGAG 2280  
 15 GGCCTGCCAC CAACATCAGC GGAGCTACTT CTTTTCAGGA CTGTGAACCC AGAGTTCAT 2340  
 GTACCTCGG ACATTTCTAC AACCAACCCA CTCACCGATG TATTGTGG CCAAGTGGGA 2400  
 CATACCAGCC TGAATTTTGA AAAAATAATT GTGTTCTCTG CCAGAGAAAT ACTACAGCTG 2460  
 ACTTTGATGG CTCCACAAC ATAACCAAGT GTAAAACAGC AAGATGTGGA GGGGAGCTGG 2520  
 GAGATTTCAC TGGGTACATT GATCCCCAA ACTACCCAGG CAATTACCCA GCCAACACCG 2580  
 AGTGTACGTG GACCATCAAC CCAGCCGCCA AGGCGCGCAT CTTGATCGTG GTCCCTGAGA 2640  
 20 TCTTCTGCC CATAGAGGAC GACTGTGGGG ACTATCTGTT GATGGGGAAA AOCCTCTCAT 2700  
 CCAATCTGTG GACAACATAT GAAACCTGCC AGACCTACGA AGGCCCATCG GACTCTACCT 2760  
 CCAAGCTAAA AACTCTGTG ATTCAGTCTA AGTCCAATA AGGGAACAGC GCTAGAGGTT 2820  
 TCCAGTCCCG ATACCTGACA TATGATGAGG ACTACAGGA ACTCATTTGA GACATAGTTC 2880  
 TGAATGGCAG CTTCTATGCA TGTGAGAACC ATCAGGAAAT ACTTAAGGAT AAGAAACTTA 2940  
 TCAAGGCTCT GTTTGATGTC CTGGGCCATC CCGAGAATA TTCAAGTAC ACAGCCACCG 3000  
 AGTCCCGAGA GATGTTTCCA AGATGTGTCA TCCGATGTCT ACGTTCCAAA GTGTCCAGGT 3060  
 TTTTGAGACC TTACAANA TGA CTAAGCCGAC GTGGCACTCA ATACAATST TCTGCTATAG 3120  
 30 GTTGTGTGGG GTTGAAGCTCT CTTCTCTCTG CATGTGACGA CAGTGGGTA TTGCTGCTCT 3180  
 CGGTATCACT GACTCATTAG AGTTCATATT TTATAGATA TACAGATATT TTGTATAATT 3240  
 GAACTTGTTT TTTCTTCCC AGCATCTGGG ATGTGAGCTG AGAATGGCTT TGAATGGCAT 3300  
 CAGCTTCTCA CTGCTGTGGG CGGATGTCTT GAGATAGACA CGGCTGTGCT GAGCTGGACT 3360  
 TTGTGACGCG TAGGTGTGAGC TCACTCTGCT TTCTGGGGTC TTACTCTCC TCAAGGAGCT 3420  
 TTTAGTGGAA AGCAAGGCC AGAATAGCT GTAAATCTGT AAACCTCAGC TTCTCTTACG 3480  
 35 CCGGGCTCTT CTAAGGGAGC CTTCTGCACT CTTGTGSCAG CTCTGACCAG GCAAGACAGG 3540  
 CAGAGGGGGA GGGAGGAGA CCGCTGCGAG CTCCTCCAC CCACCTTGG ACCTGGGAGG 3600  
 ACTCAATTTT TCCACAGCCT TCTCCAGCCT GTGTGATACA AGTTTATGCC CAGGAACCTG 3660  
 AGTCTAAGC AGTGTCTGTG AAAAAAAGAA GCAGAAAGAA TTAGAAATA ATAAAAACTA 3720  
 AGCACTCTGT GAGAGCAT

Seq ID NO: 64 Protein sequence:  
 Protein Accession #: NP\_066025.1

1 11 21 31 41 51  
 45 MGVAAGRNPQ AAWAVLILL LLLPILLLAG AVPPGRGAA GPQEDVDECA QGLDDCHADA 60  
 LQDNTPTSYK CSCKPGYQGE URQCEIDEC GNEILOGGVCH DCLNFGNYR CTCDFGMLA 120  
 HGHNCLDLD ECLNENGGCC HTCVNMGSY ECCCCKEGLF SDNQHTCHIR SEBGLSKMRK 180  
 DHGCSHICKE AFRGSVACEK RGFPELAKNQ RDCILTCNHG NGGQSHSCDD TADGPSCHC 240  
 PNYKMHITDR SKLIRJEDVT EVTNSVNTYR VLDKIRKVR LLMETCAVN GGGDRTCTD 300  
 STGVHSCSVF GPTLQDQGT CKDIDBQCTR NGGCDHECKN IVGSFDCGCK KGFLLITDEK 360  
 SCQDVEDCSL DRTCDHSCIN HPGTACACN RGYTLVGFH CGDTNBSIN NGGCGQVCVN 420  
 TVGYSBOQCH PGYKLHWNKK DCVEVIGLIP TSVSPRVSLH CGKSGGGDGC FLRHSIHL 480  
 SSDVTITRTS VTFLKNEGCK SLKNABLPFR GLRPALPEKH SSVKPSFVY NLTCSSQKV 540  
 55 PQAPRPSSTP KEMELTVEYF LETNKGRTYA SCLSKNRYK TEKLKJKAIR TLKAVKATP 600  
 FHLQLSQML DVAKPKPRTS BRQESGCVQ GHBAENQCVS CRAGYTDSA KENCILPNP 660  
 TPNBEGQMT CEPFRGNS GALKTPAWN MSECQGLQCP GBYSADGFAP CQLCALGTFO 720  
 PEAGRTSCFP CGGLATKHQ GATSPQDCBT RVQCSPHFY NTTTHLCIRK PGVTYQPEB 780  
 KNNVCSPGN TTDPDGSNT ITQCKNRCOG GLEQDPTGY ESFNYFGNYP ANTECTWIN 840  
 60 PPFKRRILIV VPHILPHED DGGDYLVRK TSSNSVITY ETQCTYERT AFTRSKKLW 900  
 KHKSNBNS ARQPLVYPTT VEDYQBLE DVNRQBLEA SENHQEILK KKLILALFDV 960  
 LAHPQNYFRY TAQESRUMFP RSFIRLLRSK VSRFLRPLYK

Seq ID NO: 65 DNA sequence  
 Nucleic Acid Accession #: NM\_007210  
 Coding sequence: 1-1869 (underlined sequences correspond to start and stop codons)

65 1 11 21 31 41 51  
 70 ATGAGGCTCC TCCGACGAGC CCACATGCCC CTGGCGCTGG CCATGTTGGG CTGCGCCTTT 60  
 GTGCTCTTCC TCTTCTCTCT GCATAGGGAT GTAGACAGA GAGAAGAGC CACAGAGAAG 120  
 CCGTGGCTGA ATGCTAGATC AATGCCCAAG CTCAAAATCA GGGCTCCAGA AGCCAGACAG 180  
 ATGAACAACC TTGACGATTC AATGCCCAAG CTCAAAATCA GGGCTCCAGA AGCCAGACAG 240  
 ACTCTGTCTT CCATAAACA GTCCGTGCTC CATTGGTTCT ATACCCACAG TGAACATGAG 300  
 75 CCGCTCTGGG AAGCGGCCAC ACAGGACCCC AATGCCCTGT GGGCAGATGG AAAAGCATTT 360  
 CAGAAGAGCA AGTGGACCCC CCGTAGAGAC CAGGAAGAG GAGAAAGGCTA TAAGAAGCAC 420  
 TGTTTCTATG CCGGATGATC TCCCTGCGAA GTTCTCTGGG GCCACACACC 480  
 CGACCACTGT AGTGTGTGGA CCAGAAGTCT CCGGCTCTCC CCGCACTGGC CACCACACAG 540

GTGATCATTTG TGTTCACAA CGAAGCCTGG TCCACACTGC TTGGAAACAGT GTACAGCGTC 600  
 CTACACACCA CCCCTGCCAT CTGTCTCAAG GAGATCATAC TTGTTGATGA TGCCAGCACCA 650  
 GAGGAGCACC TAAAGGAGAA GCTGGAGCAG TACGTGAAGC AGCTGCAGGT GTGAGGGTGT 720  
 GTGCGCAGG AGGAGCGGAA GGGGTTGATC ACCGCCCGCG TCGTGGGGCG CAGCGTGCCA 780  
 CAGGCGGAGG TGCCTACGTT CTGGATGCC CACTGTGAGT GCTTCCAGGG CTGGCTGGAG 840  
 CCGCTCTCGT CTGAATCTGT TGA GGCACAG ACAGGTGGTG TGGCGCCAA CATGTGTACC 900  
 ATGAGCCTTA ATACTTTTGA GTTCCGCAAG CCGCTGCCAA GGGGACAGAT CCATAGCCGA 960  
 GGCACACTTG ACTGGAGCCT GACCTTCGCG TGGGAAACAC TTCTCCACA TGGAAGCAG 1020  
 AGGGCGAAG ATGAAACATA CCCCATAAAA TCCCGCAGGT TTGCTGGTGG CTTCTTCTCT 1080  
 ATCCCAAGT CTAATCTTGA GCACATCGGT ACCTATGATA ATCAGATGGA GATCTGGGGA 1140  
 GGGGAGAAAG TGGAAATGTC CTTCGCTGTG TGGCAGTGTG GGGGCGAGT GAGATGATC 1200  
 CCGTCTCTGT TGGTAGGCA TTGTTCCGCG ACCAGAGGCG CCGACACTT CCGGAGAGGC 1260  
 ACTAGTGTC TTTCTGCAA TCAAGTGCOC CTGGCGAGAG TCTGGATGGA CAGTCAAGAG 1320  
 AAGATTTTCT ATGAGAGAAA TGTCCAGGCA GCAAAGATGG CCGAAGGAAA ATCCTTCGGT 1380  
 GACATTTGCG AACGACTGCA GCTGAGGGAA CAACTGCAC TGTCAAACTT TTCTGTGATC 1440  
 CTGCAAGAAT TCTACCAAGA GATGTTTGT CTGAGCTGTA CCGCCACCTT CTATGGTGCC 1500  
 ATCAAAGAC TGTGACACCA CCAATGCTCT GATGTGGGCG AGACAAAGC GCGGGGGAAG 1560  
 CCGCTCATCA TGTACTCCTG CCACGGCCTT GCGCGAACCC AGTACTTGA GTACACACT 1620  
 CAGAGGGACC TTGCGACAAA CATCGCAAG CAGCTGTGTC TACATGTGAC CAAGGGTGCT 1680  
 CTGGGCGTTG GGAGCTGTCA CTTCAGTGC AAGAATAGCC AGGTCCCAA GGAGGAGAA 1740  
 TGGGAATTGG CCGAGGATCA GCTCATCAGG AACTCAGGAT CTGGTACCTG CTTGCAATCC 1800  
 CAGGACAAA AGCCAGCAT GGCCTCTGCA AATCCAGTG ACCCCATCA GTTGTGGCTC 1860  
 TTGTCTAGG ACCCAGATCA TCCCGAGAGA GAGCCGCCAC AGCTCTTCA GGAACAGGA 1920  
 TTGCTGATGT CTGGGAACCT GATCACCAGC TTCTCTGGAG GCGTAAGAA TGGATTCTTA 1980  
 AAOCCACTGG GTGGCAAGCC AGGACCTTCC TAATCCTTGC AACACATTTG GGCCCATTT 2040  
 CTTTCTTCA CACCGATGGA AGAGACCAT T AGGACATATA TTAGGCTAG CGTTTCTGCT 2100  
 TTCTAGAAAT AGAGGCTCCC AAATAGGAGA AGGCACGCTG GGGAGGGTCT AGGACACAA 2160  
 TCTGTAGTG AAGAAGATA CTTCAGCTG GGCACAGTGG CTCATGCTCT AATCTCTAG 2220  
 ACTTTGGAAA GACAAATGTTG GAGAATGGCT TGAGGCCAGG AGTTCAAGAC CGGCTGTAGC 2280  
 AACATAGTGA GGAATCCACT TCTACGCCCA CCGTCCCGCC GGCACAAAAA AAAGCTGGGT 2340  
 ATGGTGGCTT ATGCTGTGAG TCGCAGCTAC TCAGAAGGCT GAGGTGGGAG GATTCCTTGT 2400  
 TCCCGGAGG TGGAGCTAC AGTGAAGCTT GATTGTTGTA CTGCACTCA GCTCTGGCAA 2460  
 CAGGTAGAC TGTCTGAAA ACAAAGAAA AAAAGAGAAA AGAAAGTAC TTCTACAGC 2520  
 ATGTCTATT CTTGTGATC CCAAGCACC TGCAGATGCC AGTGAATGA TATATTCTGG 2580  
 CTGGGCACAG TGGCTCACAC CTGTAACTCT AGCACTTTGG GAGGCGAAG CAGGTGAGT 2640  
 ACCTGAGGTC AGAAGTTTGA AACCAAGCTG GACTACATGG TGAACATCCA TCTCTACTAA 2700  
 AAGTACAAAA ATTAGCTGGG CATGATGGCA CGCACTGCA GTCCAGCTA CTGGGAGGCG 2760  
 TGGGACAGA GAATCACTGG AACCCAGAGG CCGAGAGTTG CAGTGAACCA AAGACGACC 2820  
 ATTGACGCC AGCTTGAGCA ACAGAGACCA AACTCATCT CAGGAAAAA AAAAAAAA 2880  
 A

Seq ID NO: 66 Protein sequence  
 Protein Accession #: NP\_009141

1 11 21 31 41 51  
 MRLLRRRHMP LRLAMVGCAP VLFLPLLRHD VSSREFATEK FWLKSLSVRK DRVLDLMLBA 60  
 MNRLRDSMPK LQRAFEACQ TLPINQSL PQFYTPAEK FFWERPQDP NAFADGKAF 120  
 QRSKWTLPT QRLKATKIDN IAFASDRI SLQSKGHT EPVCEQDQK RKPCTTIT 180  
 VIVFHNHAW STPLTFTVSV LHPTFAILL ENLVDDAST EHLKIKLIR YVKQLQVRY 240  
 VKQHEERKGLI TARLLGASVA QAEVLPLDA HECFCHGWLH PILARIEDK TVVSPDIVT 300  
 IDLNTFEFAK PVQRGRVHSR GNFDSLPTG WETLPHEKQ RKDEDTYPIK SPTFAGGLFS 360  
 IPKSYFHHG TYDNQMEIWG GENVIMSPRV WQCGQLLEH PCSVGVHVR TKSPHTFPKG 420  
 TSVIARNQVR LAEVMWDSYK KIFYRNLOA ARMAQEKSG DISERLQRE GLGHNFESWY 480  
 LIRVYTFPV PDLTFTGKA ENLTGNQL DVGNENRGL PLMYSKHL GKNQYFEYTT 540  
 QRLDRHNIK QLCLFHSKGA LGLGSCFTPT KNSQVPKDEE WELAQDLIR NSGSGTCLTS 600  
 QDKKPMAMP NFSDPHQLWL FV

Seq ID NO: 67 DNA sequence  
 Nucleic Acid Accession #: NM\_014112  
 Coding sequence: 600-4484 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 TTCTTCGCG AGGCTCCTT TGATATTAA AGTTTGTGTT TCTTGAACCT GACGATAATC 60  
 GGGGAGACTG AGGTCTCGAC AAGCGATAAC ATTTCTGATA AAGACCGAT CTATCTGCAA 120  
 TCTCTAGCGT CCTCTTTTTT GG7GTCGCG GTTCTCCAG ACCTCGCGTC CTTCTGATTG 180  
 CTCTCTGCGC TTCTATTTC TTTTTTTTT TTTTAAACAA AAAACAACAC CCGCTCCCTT 240  
 CTCCACCGCG GCACCGGCGCA CATCCTTGCT GTATTCTCTT TCTCTTCTC TCTCTCTCTC 300  
 TCTCTTTTT AATAAGAGTG GGGGAGGAAA AGGGGGGGGA GGCAGGAAAG CTTCTTTTTT 360  
 CCGTCCCGCG CATTAATGCA AGTACACTC TGTAAACAC ACAGAAAGCTG GCGCTCTCT 420  
 GCGCGCGCG CCAACACTT TCGGCTGCGC GAGGCTGTTC TTGACGATTA ATCAACAGAT 480  
 GTACAGATCA GCTCTCAAAA TGTCTCTGT GTCTCTGAG GGTCTCTTAA GACAATTGCA 540  
 TTAGCTTCT GTGATGTGAC TAATAGAAAT AATAATTGTA AAAAGACAC TAAAGCGACA 600  
 TGCTTATGA AGTCAATGCT GGGTATGATT TTACAAATAT GTTCCGGAAA AAGAACCCCC 660  
 CTTGAGAAA CTTGAGATCT GAAGGCGAG GCGGACACTT GAGGCTATA GGTACAGAA 720  
 CGAAGTATC TGGAAAGAC AAAGAATTCT CTGCAGATCA GATGTGAGAA AATACGATC 780

5 AGAGTGATGC TGCAGAACTA AATCATAAGG AGGAACATAG CTTCGATGTT CAAGATCCAT 840  
 CTCTAGACAG TAAGAAGGAC TTGAAAAGCG CAGTCTCTAG TGAGAAAGGCT GGCTTCAATT 900  
 ATGAAAGGCC CAGTAAAGGGA GGAAACCTTC CCTCCTTTCC GCATGATGAG GTGACAGACA 960  
 10 GAAATATGTT GGCTTTCTCA TTTCACAGCTG CTGGGGGAGAT CTGTGAGGCC TTGAAGTCTC 1020  
 CCACAAGAGC AGAGCAAGAT GACCACTTAC ATATGTCGCTC CACCCCTTCA GGGAGCTCAC 1080  
 TGAGAGCAA GGAAGATCAC AAGATGTGAC CAAAGGCTAC AGAGGAACCA GGGCAGACAC 1140  
 AGAGTGTGCA AGCCAAATGT CAAGGTTTGA GCCCACTTTC AGTGGCCCTCA AAAAACCCAC 1200  
 AAGTGCCCTC AGATGGGGGT TGAAGACTGA ATAAATCCAA AACTGACTTA CTGTGGAATG 1260  
 15 AACACCCAGA CCCCGGCACT CTGTCTCCAG AGCTTCAGGA CTTTAAATGC AATATCTGTG 1320  
 GATATGTGTA CTACGGCAAC GACCCCAAG ATCTGATTA A GCACTGTCCA AAGTATCAAC 1380  
 TAGAGCTCA TAACAGTACC AGGCAAGATG CTBAGTGCGA CAGCAAAAT TTGGGCGCT 1440  
 ATAACTAGT GCACTTCCAG CATTCCAAAG ACTTCAGAAA GGTCAACCGT TCTGTGTTTT 1500  
 CTGGTGCTCT GAGGAGATCA AATTCTTCAA GGCCCTGTTT ACTAAATGGG ACCTATGATG 1560  
 20 TCGAGGTGAC TTGAGGTGGA ACATTCAATT GCATTGGAAG GAAAACACCA GATTGCCAAG 1620  
 GGAACACCAA GTATTTCCGC TTGTAATTTCT GCAATTTTAC TTATATGGGC AACCTATCCA 1680  
 CCAGATTAGA ACAACATTTT CTTCGATCTC ACCCAACAAA ATTAAGATCT TCTCTTCCCT 1740  
 CTTGTAGTGT TCCAAAACCT TCAGAGAAAA ACTCTACAAA GTTAACTCCT TCCCTTCAAT 1800  
 25 CCAGTGATTC TGGAGACTTG GGAAAATGGC AGGCAAGAT AACAGTCAAA GACGAGAGATG 1860  
 ACATCTCTGT TGGGTACTCA GTGCCATAAA AGGCCCTTGA TTCTCTTAGA CAAATAGTTA 1920  
 CAGAGGCCAC CAGTTACTAC TGGTGTAAT TTGTGATGTT CAGCTGTGAG TCATCTAGCT 1980  
 CACTTAACT GCTAGAACAT TATGGCAAGC AGCAACGAGC AGTCAGATCA GCGCGGCTCT 2040  
 ATCGAGTGT AATAGTAAAG CTTCGAGGCG GCTCTGTCTT TAATCAGAA TATTATGCCA 2100  
 AAGTTTCAGA AGGAGAGACA ATGACCAAGA CAGACAAGG CTCGAGTGGG GTCTAAAGAAG 2160  
 30 AGGACTTCTC CAGCAAGGGA GCCGAGGATA ATATGGTAA GAGCTATAAT TGTCACTTCT 2220  
 TGACTTCCG ATATTCCAAA AGCCATGGCC CTGATGTAAT TGTATGGGG CCACCTTCTC 2280  
 GTTACTATCA CAGCTCCAT AACATTGACA AGTGTACCAT TAACAGCTGT CCATCTCTGC 2340  
 CAGAGAGACT TTGACATCA GAAACAAGAG CTGAGAGAA TACTATTCG TTCTGTCTTA 2400  
 35 GAAAGATRA TTGTGCCAC TTGTGCCACT TGCTCTGTCA CTGTGCTC GTGGCGGCTB 2460  
 GAAGCTCCCG AGTCAAACT CAGTGCCATC AGTGTTCATT CACCACCCCT GAGCTAGATG 2520  
 TACTCTCTT TCACTATGAA AGTGTGACTG AGTCCCAAGC ATCGGATGTC AAACAAGAAG 2580  
 CAAATCACTT GCAAGGATGT GATGGGCAGC AGTGTCTCAA GGAAAGCAA GAACACTCAT 2640  
 GTACCAAAAT GATATTTT ACCAATGTGG AAGAAGATCT TTCCGACAC TACGAGAGAG 2700  
 40 CACACAGTGT CTACAAATGC CECTGACTGA GTTTTACAGC TGCCGATACT GASTACTAC 2760  
 TGGAGCACTT CAACACTGTT CACTGCCAGG AACGAGACAT CACTACAGCC AAGCGGAGAT 2820  
 AGGAGGGTCA TGCATATCC ACCATCAAAAG AGGAGGCCAA AATTGACTTC AGGGTCTCA 2880  
 ATCTCGTAAC TCCAGACTCT AAAATGGGAG AGGCAATTTT TGAGAGTGTG GTGGAAGAGAG 2940  
 AAGAGTGGGA AGGAGAAGAC GGGCTCAAG AGCAAAATTT GACCGAGAT TCTACTGATG 3000  
 45 ACCCTCCGA TTGACTCTTG AGTGGCGGAG ACATCTGCG GAGCTGTGCT TCAATACCC 3060  
 AAGCAAGCTT GGGGCTCTGT ACCCTGTGTT CTGGCACCCA AGAGCAGACA AAGACTCTAA 3120  
 GGGATAGTCC CAATGTGGAG GCGGCCCATC TGCGCGAACC TATTATGGC TTGGCTGTGG 3180  
 AAACCAAGGG ATTOCTGAG GGGGCGCCAG CTGGCGGAGA GAAGCTGTGG GCGCTCCGCC 3240  
 AGCAGATATC TGCATGCGGA GAAACAAGT CCAAGGATGA ATCCAGTCTC CTGTATCGGA 3300  
 50 GCGCTAGAGG CTGCGTGTG TTGTGTGCA ATTOCTGAG CACAAAGACC TCTCTTGCG 3360  
 GAAGAATCC AAATGTGCGGA TATGTATGCA ACGGTGTGTT CCTCTACCAG AAGCTTCACT 3420  
 CGACTCCAG GCTTTAAAC ATCATTAAC AAAACAACGG TGAGCAGATT ATTAGGAGGA 3480  
 GAACAAGAAA GCGCCTTAAC CCAGAGGACAT TTGAGCTGA CGAGCTCAAC AAACAGCAGA 3540  
 GGGGCAAGCA TGAGGAGCAA GTCAATGGAA GCGCGTTAGA GAGGAGGTCA GAAGATCATC 3600  
 55 TACTGTAAAG TCACCAAGGA GAAATTCAC TCCCAAGCTT AAGTAAATAC GAACCCAGG 3660  
 GTTATATGCA TAAAGAAC CTCTCTCAGC AGGAGCTCTT GGTCAACAAA ACTGTGATGA 3720  
 TTCAACAAAG TCGCAACCT TTGCACATC AGATAAAAAG TCTCAGGAA AGTACTGGAG 3780  
 ATCCAGGAAA TAGTTCACT GTATCTGAAG GGAAGGAAG TTCTGAGAGA GGCAGTCTCA 3840  
 TAGAAAAGTA CATGAGACT GCGAAAACCC CAAATATTTC ACCACAGGC AGGCCATTG 3900  
 60 AAAAGVACCA GTACCCACTT TTGAGCTTC CTTTGTGTA TAATGACTCT CAGAGTGAAG 3960  
 CBAATGCTT GCGTGTCTGG AGTAACTATA AGCTCTGAT TCTGCGGAT CTCCACTCT 4020  
 TGAGTCACTT GCGTGTGCTTA CAAATATCTT GCCAAAACCTA TGTGCTTAT CCGACTCTCA 4080  
 65 ATCTGCTCTC TCATTTTCA GCTGTGGAT CAGACAATGA CATTCCTTA GATTTGGGA 4140  
 TCAAGCATTC CAGAGCTGG CCAACTGGA ACUGTGGCTC CAAGGAGAA ACAGAAAGGAC 4200  
 CACCAAATGT AAAAATGAAA GGTCCCTTGA ATGTGATGAA AACAGAGAAA TTGTATAGAA 4260  
 GTACTACA GA TGAACCTTGA CAAAAATGTG TGACCTGTGG GATCTGCTT CTGTGATGAG 4320  
 70 TGATATGTT CTTTGTATG ATAGTCTTAC AGTGCCTATG GTCTGCTAG TGAGCAGATAT 4380  
 CGCAGCATCT TTGCAGGAC AAAATATGACT TCACAACACA TATCCAGAGG GGGCTGTGTA 4440  
 GGAACAATGC ACAAGTGGAA AAAAATGGAA AAOCTTAAAGA GTAAAACCTT AGCACTTACG 4500  
 ACAATATAAT AAAAATAGGT TTCTTGATG GGAATTCAT AGCTGTAAAT GTCTATGAA 4560  
 GACCTATTAA AAAAATAGT CATAGAGCCT GCGTTATCCA AGATGAGA ATT CCGTCTTCTT 4620  
 75 GTATGTTT CTTTGTATG ATAGTCTTAC AGATTAATAA AGTGAAGATA ATGTGCTATG 4680  
 AGAAGAAATG GAAGATGTGA AACAATCACT TTTTAAACCT TGTAAGTCA AAACCACTCT 4740  
 GCGTAATATG TACTGGGGAA ATAAATCCATA AGAGATATCA CCAGACTAGA ATTAATATAT 4800  
 TTATAAGAAA AGAGAOCCAA ACTGTCTAGA ATTTGAAGGT GTTACATAT TATTACTATA 4860  
 AAGGACTACT GAGCTGCCCA TTGGACCAT TTGTCCAAA CCGATAAAT TTGTGCTTAA 4920  
 80 TTATAATGA TAAAGAAC CTAGCAGAGG GAGGAGAAAT TGAAGTCCA GGGCAATGA 4980  
 AAAAATATG CCGCCTCTCA ATTAGTCTT CHTCATTTG CAGTGTTC AATTGTTGAC 5040  
 TAGAAAATGG AGCTGTGTTT AGGCTTGTGT AGAGTGACAG AAGCAACATG ACAGATGTGT 5100  
 GCAAGCTGTT TTACCCAGC CCGTGGCTTGA CATACACATG CACACCTCT CTGATATTT 5160  
 TTGCTCTTAG ATGTTCAAAT ACTCAGATG CTCTTTGTT GCGGTTTGA TCAATTTGT 5220  
 85 CCACAATG CAAATATG TCCCGAGGCT TCTGTATGT ATCTGATTT 5280  
 AGCCGATAT TCTCTCTG TTGTGTGTA ACCAGATATG ATTTCCTT CTAAGCTCC 5340  
 TTGTGGTAC TAATCTCACT TGGCACATTA TAACATAAGG AATCCCTCTA ATCTAAAGC 5400

ATAGATGGAT ACAAAATGTCA GACCGTGGGT TTAATTGTGT TAGAACACAT GGCATTCTCT 5460  
 CACAAGGTAA CCGTCGTGTAT TTAATTATTT TCTTTGGGT AAATATTAAT TCCAAACCTT 5520  
 GTTGTCAGGC AGCGCTTAA GGTAGGTAC CACAGACTGA CAOTGTGTAT ATGTACCAAC 5580  
 CAATCCCTTC ATTAATATGTACAGATTITA GTTAGTAGAG ATTAATATAGG ATTCTTAGAA 5640  
 5 GTATGTCCTC ATAGAACTTT TAATACITAA GCGTTGTGAA AANAATATCCA TGAAGGAAAA 5700  
 GCTCCTCAGC ATACCTGCTC AGGGAATATG GCGTAAATAA CTGAACATTA AATATATGTG 5760  
 TAAAGTGCTC GTTAGTGTAG CCTCAATGCT TGCTACAAGG ATGTATATGAC AAGAGCTGAC 5820  
 TTTAAATAT TGTGATATAT TGTCCAAAC AGTAGTTAT TTTTGTCAC GGAATGTGAG 5880  
 AAGATATTAC AAGCTACTGTC ATGCACGTGC AGATTAACIT ATTTCAATTA AGAAGTTGGG 5940  
 10 AGAACAAATA GGAIAAAAAA AACITATTTT TCTAGTAAT ATTAATGTAT TACATTTTCA 6000  
 TTCCATCATG CATTAGAGTC AGTTGTGBCCT CCGCTAGCT GTTTACATTT GCAAAATGTA 6060  
 CAACAAGGT AATAGAGCA CTTATTTCT TGCAGTAGAT ATCTCTTTGT GTGTGTGTGT 6120  
 GTGCATTTAA GTTGTAAACG GTAACTGAAA ACAAAATGAAA GTTCTTGCTA TAATGTGTATG 6180  
 15 GAAAACAAGA AGGAATAAGA AATATTTTTA TGCCTACTTA GGAIAAAAAA GGTAGCACCT 6240  
 ATTCATTTCA AGTACTTTTT TTTTTTAAT TTTTAAGCTC TTAAGCTACA TIGTTATGCT 6300  
 TAAGATGATA AACATAATAC CTCITTTTAT TCGTTTGTCT ATGTTTCATA TGAACACTTT 6360  
 CAGAAATAT TTTGATAAGT GTTGTGTGAA TGTGCAACCG TGAATTTTTT TGTCACTCTG 6420  
 TAGTGGCAT TGCATTCAT TTTTACATA ATTGCCAGTT GCTTTGTATC ATTGTTTFTGT 6480  
 20 TTGGGTTTTG TTTCTTTTC ACAGTGCCTG GTCTTGCTTT CTAAAGTTGT GATGGCAGCT 6540  
 AGAGTTCAAC CAGTTCGTGA CTGTGTGAGC GAATGAAGTT AAAAAAATGT CTTCTTGATG 6600  
 TTGTGTTTCT ATTTTCAITT TTGCATTTTT TTGTTTGCT ATTTAAAAAA GAGAANAAGG 6720  
 TAAGCAGAG ACAGAAATCA GCACTAGTGT CTTGCTGTCA GTTTTCACTT TAAGCGCTCT 6780  
 25 TTCTGTGAT TGCACGTGTG CCGAGCTCTA ATATTCACAT AAGCTGAAT AAGCTGTGCG 6840  
 AATGGAAGAG TTGACATCTC AAATATGTG GATGAATTTA TCTCTCTAG GAATTTTGT 6900  
 GGATGTCATCT CAATATGTAT AGCCAGACTT GAGAGGTGAC AATTAAGAT CTAIAAAAAA 6960  
 GAGGAGATTC CCCCACACAA CAATATTTAA TTTCTCTAT AAAAAAGATA ACAGAACATCA 7020  
 TGTGGGCAAT CTTAAGCAAA CATATCTAT GTGAGACTGT TAAATCAGCA AACAACAGCA 7080  
 30 AGTTTGTTTA ACTTTTGTAT TGTGCAAGAT ATTTTGTGT TGGGCAACCT ACTCTATGAA 7140  
 CAATTTCTCT AGTGTGTG GACACAATAG GTTCTTTTAT TTGTCATCT ATGCTTTTTT 7200  
 ATTTTCACTC AATTTTITTT TTTTCTGAGA CAGACATAGT AGTATCACT AGCATTTGAA 7260  
 AATACATATC ACTATCTCTG GAATATTAT GTTCACTCTA CTTTTAGTA AATATTTTTT 7320  
 GGATAGCGTT GACACGATAG ATCTATATCC ATACTTCTT ATATATGATA ATTTATTTT 7380  
 35 CATTTTTGTT TTTCATATT ATACATATTT TGTGTGAGAA GAGTGTGGGC TTTTTGAAA 7440  
 GAGACAAAAA TTTTATGATA AACTTAAGAC TCTTTTTTGT ACATATATAA GCGTTATATC 7500  
 CATCTCTCAA GATATATAT AATATTTAT TTTTAAATTT AAGATTTCTG AATATTTTAA 7560  
 TCTTAAATGT TGAATTTAAA CGAGCTATTA TGGTAGGAAA CTTTTTTAA TGAGGAATTT 7620  
 CATGATATAT TGTGATTTTT CTCTCTGGA AAGGCTTCC CTTGTGATGA AATAGTATGG 7680  
 40 CCAGCTAAAA TTGTTGCTCA TTTAAATACT GAAAATATTT TAAATATTT TGTCTATTT 7740  
 CTAAATGAG CTTTGATGTA AACTTAAGC CAGGACACAG TGAATGGCT TCAATCTTCC 7800  
 TTTTGTCACT CTTTCTCTCA TCACTACCT CTTGATTCAT TCTGTTGTT GGGATAGAAA 7860  
 AATCATAAAG AGCCAAACCA TCTCAGAAAG TTGTGGATTG AGAGAGACAC TACATGACTC 7920  
 CAAGTATATG AGAAAAGGAC AGAGCTCTAA TTGATAACTC TGTAGTTCAA AAGGAAAAGA 7980  
 45 GTATGTCCTAA TTCTCTCTAC ATGACATATT GAGATTTTTT TAAATCACT TTAAGATAG 8040  
 TGAATGTCTG TCTAAAACGT TTTCTTTTAA GTAAAGTATG ATTTTATAA ACAAGCATG 8100  
 GGGATCTTT TCTAAGGTAA TATTAATGAG AAGGGAACAA AGTATCTTA ACAGTTCTTT 8160  
 GTTGAAGCTT GTGTAGACAC ATTAGTTTAA TAATGTGACA TGTGCACATA TCTATTATGT 8220  
 ATCCAATGCA AATCAGCTC CAATAATATT AATATATAT ATATTTTTAA ATGCTGAGG 8280  
 50 AATATCATTT TTTCTATATA ACTGAAGAAT CTCAGTATGG CTATTAATAA AATATATAGC 8340  
 CTCCTGTGT GTGTGCTGAAA AACATACAAA AGTGAACCGT CTTGAGACTT GAGACTCTT 8400  
 GCGCTTTTGA GTTAAATATG TTAAGTCACTT TCTAGAGGGG GATCTTCACTG CATTGAGGTG 8460  
 TGGAAATTTG GACTGAAGAA CGTGTGTGCT TGTCTTGTTG CTGTAAGCCA CCGTATTGCT 8520  
 TTAAGTTGAG AGGAGGTCAA CTTTAGCTAC TGCTTTTGTT TTGAGAGCCA TGGCAAAAAA 8580  
 55 AAAAAAAGAA AAAAAAGATCA AGTCGTCTT GGTGAGCCAG TAAGGTGAAA CTTGCTGACG 8640  
 TTGTCCAGGC ACAAGAGAAA ATTAGGAAT TGAATGCAA CTTGAGTATC AACTAATATA 8700  
 TTTCTATCAA AGTATAGTAC TTTTGTGTTG ATTTCTTAAA CGAGGCATCT GCAATATAGA 8760  
 AGAAGATAGA AGACAGGCGG TCGGAGCTTT GGAGGCGCAT GTTATTTTCC CAAGAAGA 8820  
 CGGCAAGGG CAGAGGCATG GATTCCTTTC AGAGCACTTC CTTTGTGTT TCAAGTATCT 8880  
 60 TTTCATAGAC AGTGGGCTCA CATGTTCTGT ATAGTGTCTG AGTTGCTIAG AAAGCATCC 8940  
 AGTTAATGCG AGTAAATAGA ACTTCTGGAA TATGCTAGGG CAGAGATGAT TCAAGTATGT 9000  
 CACATGAGAA AATGTGTGAA TCAACAGATA ATACACAGT GAGAAGACTG ACATATGATA 9060  
 TTCAATGTCT CTTGTAATA GAAAGAGAGA GGTGATGACT TCACTGTGTC CTACACCGGA 9120  
 GAAAAGCAGG AATAACTTAA CCGTGAATAT AATGTTTAGT TTTTATCAGA GAAAATGTGT 9180  
 65 TCTTAGAGC ATAGAGTCCC AAAACTCAAT TCTGGTTTTC CCGTGTTTTT TTTTTTTTTT 9240  
 CTTTCCCAAT TGAATAATAT CAGCATATCT ACTTTTTCTT TTTGTGCTC AGGTTCTCTA 9300  
 CTTGTAAAAA AATGTGATA ATGATTAAT AATATATTA ATAATAATA TGTATATGA 9360  
 GTACTTGTGT GTAAAGCAT CTGAGATCTT TGTGTGAAAG GCACATAGG AGTGGCAAT 9420  
 ATATTATGT GSCCAAGGGG GTTATTTAAA CTGTGAGTTC CCAAGGCCA GGAAGGTTGT 9480  
 70 GGGTCAATTT TCTTAAAGC GAGCTGTAAA TATCAACTAG GAGCACTATA TGTGTGATA 9540  
 TGAAGATGCA AAACATATTAC TAGGCTGATA AAATCAATAT TTTCTAATGG CTACCAATAA 9600  
 GGCATAATAT ACAATAATAA AGCCAAATTT CTTTAGGGC GACTATTGA CAACCAATG 9660  
 GAAAACCTTG GGGGAGGCTT GAGGTGAAA CATCTCAATA TGGCAATGTA AATATTAAT 9720  
 TACATCAATA TCTTACACCA GAAAATGTCT TCTATATGGA ATGATTTTAT GTTGTAAAGA 9780  
 AAGAAGATTC AATTTTGTAT CCGATTTGTA ATACTAGAAT GTTGGCTATA ATAGTTTCT 9840  
 75 TCTTACAACA CATGAAATTT TTTGTTTTA TTTATTTTG TTTTATATG GCATGTTCT 9900  
 TCTTACTAC AAACATGTT TGTGTTTTT TCTATGAAA ACAATCTTCA GGCAGCAAG 9960  
 ATGTCTGTTA CATCTAACT TGAATAAAT AGTTTACCA CCAATTTAC A

## Seq ID NO: 68 Protein sequence:

Protein Accession #: NP\_054831

5  
1 11 21 31 41 51  
MPYEVNAGDY FTMNVKKNP PLRNVASGIE GQLEHFGTE SKVSGKNKF SADQMSBTD 60  
QSDAELNKH EELHVDQD SSSKSLDLES AVLSEKAGN VESPKGNF PFFIDEVTD 120  
10 RNMLASPTA AGGVCEPKKS PQSAEADDQ DMACTSGDS LETKEDQKMS FKATEETQA 180  
QSGQANCOGL SPVSVASKNP QVPSDGGVRL NKSXDTLLVN DNPDPAPLS ELQDFKCNZ 240  
GYGYVNDPT DLDHPRKHY LGLHNKTRQD AELDSKILAL HNMVQFSHSK DFQKVNRSVF 300  
SGVLQDINSR PVLNNGTYD VQVTSGGTF GIGRKTPDQ GNTKYFRCKF CNFTYMGSS 360  
TELEHFLQT HPNKAISLP SSEVAKPSBK NSNKSPALQ SSSSDLGKW QDKITVAGD 420  
DTPGVSVPL KPLDSRQNO TEATSYVYVCK FSPSCSSS ELKLEHYVK QHGAQVSGGL 480  
15 NPHLNDKLSR GSVINQNDLA KSSEGETMTK TDKSSGAKK KDFSSKGAED NMVTSYNCQF 540  
CDFRYSKSHG PDVIVVGLL RHVYQLINBH KCTKHCPFC PRGLCSPEKH LGEITYPFAC 600  
RKSNCSHAL LLLHSPGAA GSSRVKHQCH QCSFTTDPVD VLLFHYESVH ESQASVQKQ 660  
ANHLQSDGG QSVLESKSKS CTCKDFTQV EBSIRHYRR AHSYCKRQK SPTAADTQSL 720  
LSEHNTVQIQ EDDITANOE EDGHAISTK EEPKIDERYV NLLTDSKMG EPVSESVK 780  
20 EKLEEDGLK EKVVTSSSD DLRNVTWEGA DILRGSPSYT QASGLLTPV SGTQEQITL 840  
RDSFNVAAH LARPTVGLAV ETGQFLQAP AGGKSGALP QVYPASGHNK SKDBSQSLLR 900  
RRRSGVFCA NCLTKTSLW RKNANGGVVC NACGLYQKLH STPLNLIK QNNGEQIR 960  
KTRKRLNFEA LQAIBQNKQQ RGSNEBQVNG SPLEKRSDEH LTESHQREH LPGLSKYEAQ 1020  
25 SLKTSISIAQ QVLYVSQTLD IHKMAQPLH QKSPQSTQ DPNNSSVSG GKSSIERQSP 1080  
IEKYMSPAKI PNYSPGSSA EKYQPLRGL PFVNDPQSE ADWLRFWSKY KLSVPGNHFY 1140  
LSHVGLPNE CONVYYPYPT NLPFFSAVG SDNDIPLLA IKHSRFGPTA NGASKERTKA 1200  
PPNVKNBPL NVVKTQVDR STQDELSTK VHCQIVLDE VMYALHMSCH GDSGFFQCS 1260  
CQHLCTDKYD FTHIQRLGH RNAQVBEKNG KPKE

## Seq ID NO: 69 DNA sequence

Nucleic Acid Accession #: XM\_073879

Coding sequence: 1-387 (underlined sequences correspond to start and stop codons)

35  
1 11 21 31 41 51  
ATGGGGTTTG GAGACACGGG AACGGTGAA GGGAGCCTAG GAACCTCGAA AAAACCACT 60  
GAAGTGA AAA TGGTTGAGC CAGTCAAGGT TTGCTGACAA TGGAAACAAA CCAGTCCGTG 120  
GCACAAAGCA CAGGCTGCTC AGTAGTAAAG GTAGACACTG TCTCTGTGCA GAGTITATAC 180  
CACTGGCGCT TGAACAAGG GAGCGTGATG CACTGCGTTC GGGATGATCA CCCCACAGAA 240  
40 GACAGGAAG CTCACGTCTC TGCCXGAGTT GCAGCCATCG CTTCTCCAGC ACCGACTCCT 300  
GTCTGTCTG CACACGACT AACACAGAGC ATCTGCAGT TTCTACAGCA CTCAGGACG 360  
AACACTCACT TGCAGGCTGC TAACATA

## Seq ID NO: 70 Protein sequence:

Protein Accession #: XP\_073879

45  
1 11 21 31 41 51  
MGPDQGTVE GSLGTSKFP EVKMFASQD LLTMTNQLS AQGTGCSVVK VDTVLFSLEY 60  
50 HGFGEHSYM HCLGDHPQC DRKAHSPVP AAIASPAPT VCPAHSTQS ICQFLHCICQ 120  
NTHLQAAV

## Seq ID NO: 71 DNA sequence

Nucleic Acid Accession #: AB033064

Coding sequence: 826-1986 (underlined sequences correspond to start and stop codons)

55  
1 11 21 31 41 51  
GGGGAACAGC AAATTCAGTC ACAGACAATC CTCACCTGCG TCAAGAGCCA CTTTCTCTT 60  
CTCGCTTTC CCCCOCGACG GGGGTAAAGG ACTGAGCGTT TAACTTTAG CCGGTGGCT 120  
60 ACCAGTAAA ATCTACTCTA TCTTAGTTCT TAGTGGATAG CTTCCTATT TTGCCCATGT 180  
TTCTTAGAA TCCCTTTTGA ATACTTTT GTACATAGTA GTACTAGGA GTAGCAGGA 240  
GAGTGAAT AAATAGACC CTCTCTTTT CCGTTGTCAT TCAGGCCCTT TTTCTCTCT 300  
AGAGGGAAAT TACCAGTAAA CTCTCTTAAA TCTTCACACC CTCTCAGTC ATACTGTGAA 360  
65 GAAACACACT AAAGTGGAGA TTATTGACC AGTGAACAGC AACCACGCTT CAAGCATGTG 420  
TTGTGTGTGG CACATGGAGA AACACTCTTT TTAATAATC TCCCAATTAC CTTTTCACA 480  
ATGTGTATCC ACCTAGGATT TGCTCTGGG GYAAGCACT AGATTATTAT GTCAAAGCTC 540  
CCCTCTCAT GAGTGAAGC ACTGACCAAC CATGAACACT AGTAGGGAT GGGGAAAGGG 600  
GACAGAGCAG AGCCAGTGT TTCCACACTTT GGGAAAGCAG AGTAGCTTT ATCATCTTCT 660  
70 TCTGGGAGAG AGGCATAGAG ACATAAAGCT AGTGAAATG GGTGGAGGAA GAACCTCTAT 720  
TCCACAGAAC AACATGTGAA GAGAGAGAAC CAACATATA GTAAAGGAGG TAGACGTTAC 780  
ATCCAAGAGG AAAATATCCA GCGAAGGAAG CACAAGCTGA TCAAGAGTGT TAGTCTCTGTG 840  
GCTGCCAAGT TGCTGTGTTT GACAGAGCTT CCACTCAGGG AGACTATCC TCACTCTCTG 900  
ATTTTACAG CAATTAAGCG CAATATGTTT GAGGAGGAAC TGGACTTTAG TGCCTGTGGT 960  
75 ATGGATGAG TGGTGCTGAC AATGAGCAA GGAACCTTGG GTCTGCGGAT CAATGAGGAG 1020  
CTAATCACTG CCAACCCACA AGTGTTGTTA GTCAAGTAC CAACCCCTTG GGTGCAAAAT 1080  
GATAGTGACA TCACTGTTT GCCCATCTA GAGAGATGAT GATCTGGGTT AATGAACCGA 1140

CCTCAAGCCA TCCCTGAAGTC CGTTAATAAG TCTGGACAT TTCAAGAGTT GCGTCGGCAT 1200  
 GGTGTCTCTC TGGGAGATCA TTCTCTGAT GTGGGCGCAG AAAATTTTGC TAAATAGTT 1220  
 GTAGAGGCTG AAGTGTGCGA GTTCCCAATG TGAATAAAGA ATACGCGGGG TCACAGAGGT 1320  
 5 AAAGCTGTTT TCTTGGCTCG AGATAAGCAC CATTTGGGCT ATCTAAGCCA TCTTATCGC 1380  
 TCAGAAGGCG CATACCTGTT CCAGAAGTAT GTTAAAGAGT CTCAATGAAG GGAATGACT 1440  
 GTCAATTGCG TGGGAGGGCG TTGTGTTGGC ACCATGTTCAT GTTGTTCAC AGATGGGAGA 1500  
 ATGCAAGA CACTGATATT AGTGTGTGCG GGGATGATGT GTCTATGAG TGAACAAGGG 1560  
 AAGCAGCTAG CTATCCAGTT GTCTAATATC CTTGGGAGTGG ATGTGTGGTG CATTCAGCTG 1620  
 10 CTGATGAAGA ATGACGCGCT CTCTCGGCTG TGTGAGGCCA ATGCAAAATG AGTTTCTCAT 1680  
 GCGTTTGATA AGGCTGTGAA TCTAGATGTA GCTGGTATCA TAGCAGACTA TGCCGCTCC 1740  
 CTCTACCTCT CTGGCGGGCT CACCGCGCT ATGTCCGCTG TCTCCGTGGT GTCCACTGCG 1800  
 AGTGAGACTA GTGAGCGCGA GCTGGGTGCC CGACCCAGCA CTGTGTTTGA CAACTAGT 1860  
 GCAAGTTGCA GCTGTGTGA CAGCGACCT GAAAGCAGCG AGCGAGAGCT GCTCAACAAG 1920  
 15 CTCCAGGGG GCGCTGTCAA CATGAACGAG CTGTGAGCCA ATGAATCAAA ACTACTGGTG 1980  
 GACTGACTCC ACTGTGTAAT AACCAACAAA ACCCTGTGAA AACTTCTT CTCTTTCT 2040  
 ATTTTAAAA CCAACITGCA ATGCTGTCCA TGGGAGGATG TCAGGAAGAT GAGAGAATA 2100  
 TAGTAGATT AGTTGGAGAG AGTGGAGAT AGATGAGACC TGTCTAGTA AGATGTACT 2160  
 TCTATTACA TCCCTACAAA ATAGAGAGGCG AGAATAGGTG GGTGTATGAA AATGTTCAG 2220  
 CTCTCATAGT TACCCTTTTA AATTGCTAAA AAGTGTGAT GCTCATAGGC CATGAAGAAC 2280  
 20 AAATACITTT TTTTTFAT GGTCCCTGCG TTTGTFTT GTACAAAAAA AATGTGTT 2340  
 GCTACAAAAT TCCAAGTAGC ATAACTTCAC ATTTGTTTGG AAGATTGTG CATGAGGAG 2400  
 AAAACATCTG CTAAATTCAG AGGAATTTT GTATTATACA GCTTGAAAAA TCTGCTATT 2460  
 TCTTTATTA CTAGCAGCTT TAGTTTATAG TTATGCAAT CTGAGGGGCT TCTTTACTG 2520  
 25 GAGTTTCTTA TTTTFTTGT TTTTCCGCTT TAATTGGTG GGAAGTCAAA TTGAATATA 2580  
 CCAATAAAGG GCTCTTAAT GAGAAAATG GCAATTTGCG ATGATGAAAT GGAATGAAC 2640  
 AGTATTGCAA TGTCCGGTAT ACAAATAAC ATTAATTCAA TGTAGATAAA ATTACACTAG 2700  
 TTTAAAAAT GTGCATTAC TTGTATTGTT TAGTTTFTA GCTTTTGT AAAGATGCG 2760  
 TCTGTAAATG TGTCTTTT TTTTFTTT TAATACATG TAGTCAAAA TTTCCTG 2820  
 30 TATGCTGCA TGTCTAAGT GAGAAATG CACTCTTT TGTATACAG 2880  
 ACATGACTT GAAACATGTA CATTTAAGC CTTTATTT TCTCTTTT GTTTGTGAG 2940  
 TGTGGCATTT AAATAAGGAC AAGGAAAAAT ATTTTGGG GCAATCAAG AGCCTATGAG 3000  
 TCTAAGTAT AAAGCTGAAG TGAATTGAAA TCCGAGGCT ATATAITTC ATTTTACA 3060  
 TTTTACGAGG GAGTATATGT GTATGTGTG GCAACGATCG ATGTGATGT GTTTTGT 3120  
 35 TGTGTCAAT GACTAATCA AAAAGGATA TTGAGAAAAT GGAGCATGAT GGGAAACGA 3180  
 GTTTTISACT TAAATAAGA GATGAGTTGT TTTCAAGT AGACTCCAT GGGGTAGAGG 3240  
 TAITCAOCTT AAAACATAGG GTGAGTAGAT GCTTTTATG GCTTTTGT GTATATGTAC 3300  
 GTTGTGTT TTTTTCCTT TGTCTACGCG CTGTTCAGTG TACAGTTAT TCAAGGCTAC 3360  
 40 ATGCTTTCT TTAATGCTTC TGGCTATGCA TTTTCITTT TACATATAG GATTTGGAT 3420  
 TGGGGTGGG TTGATGTTT TTTTGGGG ACTATTATAG TACTATTGAG TCTCTATAG 3480  
 CCTCACITT AAGCCTTCAA TACTGTGAC CTCTTATTT GCTTACTG CAGAAITTA 3540  
 AAAAGCCCC AAATCTGATA TAATATGAGC CTTTAAACA TGGGTAAAC TAATCCACT 3600  
 GATGGGTTTG GATGTATGT TAAGAAATGG AGATGCTGCA GAGGCCAACG TAATTTT 3660  
 45 AACAGCAAGT TTTCCATCT CCTACGAATC CTCTGAAGCT TTAACCAAG CCGTTTCTG 3720  
 CCTCTCAGT GCTATTTC TCTAGATGCA CTTAACAT AATTCTTGG ACACACTAG 3780  
 AGAGACTTG AGGCAATAA AAAGAGATG TATTAACAG CTATACAGA GGTGTATCA 3840  
 50 GTTCTACTG TACAGTTT CCCCCTTT AAAAAGGAAT GTAATAAAAT TTGTTTTT 3900  
 CATAGAATTA AATAATATA AATGTAGTG AAAGGTGAT TGTGTATGAA TAGAATGTA 3960  
 CCTCTCATCT GTGCAATGTC TAAITGACC TCAGAGAAA GGATACATAA GAGGAGTTT 4020  
 TAATTATCT TAGGATATCT TAATTGATT TAAAGAACT TATCTGGC AGGTAATG 4080  
 55 GGGGACTCAC ATACATATAT TAATAGCTCT GGTCAATTA CAGAAAGAA AATCTGTAC 4140  
 TCTTCTCT GATACACCAT ACTGTGAGAG ATGCTATGTA TTGTGATAG AAGAAAAAT 4200  
 GAGGAAGAAA GAATCTGCTA ATTAATAAT CATTCATATG TCTATAGA GACCATCTG 4260  
 TTGCCATGTT TATTATGACA CATACACTTT GAATAGTAC ATATACAG TATGTAGTC 4320  
 60 ATGTTGTGT TGTGGGGTA AGGCATCAGG AAAAATGTAG TTAGTCTTT GTTAACATT 4380  
 ACCAAATTA CCAACTATAT TATAGAAAT ATGTGAAAT AGTTCAITAG CTATTATC 4440  
 TATATGAT TACTCATGATA TTAAGAATGA CACTCAGTG TTAGATGAA AAGAGAGTA 4500  
 65 GTTATCTAG ATTTAAAAA CAGGTATATC TCTTGAAAT GCTTCAAGT TGAAGTAGG 4560  
 TAAAGAGCA AATCATCTG GAGGTACTT AAGGAACAA ACAGACGAG ATATTAGGT 4620  
 TAAACTTATT TCTAATATG TTAATAACT TTTGTATAA TCTCATGCT ATTAGAGAG 4680  
 AGAATGTATA TATCAAAATAT GTGATGAT TAAAGATGAA TTGTAAAAAT TTGTATAT 4740  
 70 GTTAAAAAT TAATCTTAT TGTATTATCA TGTATTATTA TTCTGATAT TGTITATG 4800  
 TACCCAGA TGAAGACTAG ACCTTATA TCTAAACATA CAAATATGAA CTATCTATT 4860  
 TAAAAATTT AATATGTTT TCTTTTTG GTGCTATAA TGAATGGTC ATTCTGCTG 4920  
 GCTTTTCTC AATGAACATT GAAATCTTC TGTATATGT ACCAATAAGA AAATACCTC 4980  
 75 GGAACAGTAG AAAAAACCAA CAGAGACTT GGCATTATC AAGCATAA TCAGACTTGT 5040  
 AGAACATAT GAGGCGATG ACTTTGAAAA TCACTCTT TTGCAAGA GAAAGAGG 5100  
 GAGAAGCAA TTGTTTCT TTAGTGATAG CCACTTTGG GCGTTTGGG GCTTAGAGAG 5160  
 ATTTGAAAT CAAATCTGT GTTATACITT TCTCTGGCT CACTTTTT GAGAAGGTT 5220  
 ATGGGCTATT TGGCTGGTA GACACGATCC CTTCTAAGA AATGTAGT GCTCAGACG 5280  
 70 GTAACCACTG CTGCTACIT TTTAATGTT TGTGTTGTC AATTATTT AAGATTGTT 5340  
 TTTGTGATC TAGGATTITA AAAAAATGAA TATATTGAG GATTATAAC CAG

Seq ID NO: 72 Protein sequence:  
 Protein Accession #: BAA86552

1 11 21 31 41 51  
 | | | | |



MCSSVAAKLW FLTDRRRIED VPKKHLRAL KACCKEHLDEL FRAYVMDEVV LITBQXNLGL 60  
 NUNGELIAY PQVYVVPYPT PWFQSSDSIT VLRHLEKMGCG RLMNRPRQAL NCYNQFWTQ 120  
 ELAGHGVPLP DTFSYVGHEN FAKMIDEAEV LEFFMVVKNZ RGRHGRKAVPL ARDKKHLADL 180  
 SLRLHIEAPY LFQKYVKESH GRDVRVIVVG GRVVGTMRLRC STDGRMQSNC SLGGVGMMS 240  
 5 SHLRGQKQLAI QVSNILQMDY CGDILLMKDD GSPCVCHANA NVGFIAPDKA CNLDVAGHIA 300  
 DYAAASLLPSG RLRTMRSLLS VYSTASETSE PELGTPASTA VDNMSASSSS VDSDFESTER 360  
 BLTLKPLGGL FNNMQLLANE IKLLIVSD

Seq ID NO: 73 DNA sequence  
 Nucleic Acid Accession #: XM\_040080.2  
 Coding sequence: 159-1104 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CTGAAGTGGGG GCGGGGACTG CTGGAATTGC GGGGCTGCTC TGGGTAAGCG CCGGGCAGGA 60  
 CAGCTTGGAG ATAGGGGCGG GAATTGCGGG CGTCACTCTG CTCTGGGAG CTAAGCAGGC 120  
 15 TGAGGGGAGT GACACAGCAG CATTCGCGGG ACGAGAGCGA TGAATGAGAA CCGCGCACCA 180  
 GGCTGTATCT CAGACCTGAA GCTGGCTGTG CCTCGGGGCC ACATGGCAGC CAAAGGCTGG 240  
 GGTCTCTGCG AGGCGCCCTC AGTTCTCTGC CTCGACGCGT GCGTGGACAA TGGCAGCTCC 300  
 20 TTGCACAGAC TCATCCCTCT TCTCCCGCAA GACTTTTATT ACGTGGCAT GGATTTCGGA 360  
 GGTCATGGGC TCCTGTCCTA TTACAGCCCA GGTGTCCCAT ATTACCTCCA GACTTTGTG 420  
 AGTGAGATCT GAAGAGTTGT GGCAGCCCTG AAAATGGAAT GATTTCCTAT TCTGGGCGAC 480  
 AGCTTCGCTG GCGTGTGCG CGGAATCTTT CATCTATACCT TCCCGAGAT GGTGATATAA 540  
 25 CTATCTGTC TGGACAGCGC GCTCTTCTCT CTGGAATACG ATGAATATGA GAAGCTGTG 600  
 ACTTACAAGC GGAGAGCCAT AGAGCAAGTG CTGACGATAG AGGCCTCCCA GCGGCGCTGC 660  
 CAGCTGTTC A GCTCAAGCA GCTGCTGAC AGGTATCTGA AGAGCAATAG CCACTTGAGT 720  
 GAGGAGTGGC GGGAGGCTCT CCGTCAAAGA GGAACACAGA AGGTGGCCAC AGGTCTGGTT 780  
 CTGAACAGAG ACCAAGAGCT CCGCTGGGCA GAGAACAGCA TTGACTTCAT CAGCAGGGAG 840  
 CTGTGTCCCA ATTCATCAG GAAGCTGCAG GCCCATGCTC TGTGTATCAA AGCACTTCAC 900  
 30 GGATATTTTG ATTCAGACCA GAATTACTCT GAGAAGGAGT CCGTGTGCTT CATGATAGAG 960  
 ACGATGAAAT CCACCTCTCA AAGAGCAGTTC CAGTTTGTGG AAGTCCAGG CAATCACTGT 1020  
 GTCCACATGA GCGAAACCCA GCAAGTGGCC AGTATCATCA GCTCCTCTT ACAGTGCACA 1080  
 CACATGCTCC CAGCCAGCTG GTAGCTCTGG GCGTGGAACT ATGAAGACCT AGTGTGCCA 1140  
 35 GACTCAACCA TGGTATCTG GTGATGAGAG CCGCACAACA AGGTCACAAG TGGTGGGAC 1200  
 AGGCGTCACT AGCTTGGAG CCGACGCTAG GATGGTAGTC AGGGGAAGGA GCGAGATTCC 1260  
 AACTTCAACA TGTGTGACT CAAGGGGGAG ACAGAGTCTG GGTTCACAGG CTGCTTTCTC 1320  
 CTGCTCAATA ATAAATATCC AGCCAGCTGG AGGAAGGAAG GCGAGGCTGG GCCCACTTAG 1380  
 CTTTCCCTG CTGCCAACT GGATGGAAAA TAAAGGTTCT TTGATTTCTA

Seq ID NO: 74 Protein sequence:  
 Protein Accession #: XP\_040080.1

1 11 21 31 41 51  
 MSENAAPOLI SELKLAVPW HIAAKWGS LQGPFVLCIIG WLDNASSFDR LIPLLPQDFY 60  
 45 YVAMDFQGHG LSSHYPGPV YYLQTFVSEI RRVVAALKWN LRIHLQSHRG GVVGGMFFCT 120  
 PFPMVDKLLI LDTPLFLES DEMENLLTYK RRAIEHVLQV BASQPSHVF SLKQLLRLL 180  
 KSNHLSBEC GELLQRGTI KVATGLVLNR DQRLAWAINS IDPISRLCA HSIRKIQAHV 240  
 50 LLKAVHGVYF DSRQNYSEK SLSPMDTMM STLKRNQFVF EYVGRNCHVM SEFQIVASHI 300  
 SSKLQCTHML PAQL

Seq ID NO: 75 DNA sequence  
 Nucleic Acid Accession #: NM\_005794  
 Coding sequence: 434-1276 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 GGTTCCCTTC CAGCGTGTGA AGCTTTGTTT TTTTGTCCTT CATGATAAAT CTTCGTGCTG 60  
 60 CTCACTGCTT GGGCTGCTGC CACTTTAAG AGCTGTGAACA CTCACCGGA AGGCTGTGCA 120  
 CTTCACCTCT GGGGCCAGCA AGACCAAGAA TUCACCGAGA GGAATGAACA ACTCTGGAACA 180  
 CAGCATCTTT AGAGAACGGA ATACTCAACG CAAGGCTGTG CAACCTATT CTGATGACGA 240  
 GTAGGGCCAA GAACCAATCA ATTGCTGATA CATTTGGTGT ACTTTGAGGA GACTGTGACC 300  
 TATCAACCAAG TGGTGAGACT ATTGCCAAGC AGTTGAGACTA TTGCCAAGTG GTGAGAACCT 360  
 65 CACCAAGGCG TGAGACTATC AACTATGCCC AAGTGGGCTG ATTACGACAG AAGCATCTCA 420  
 GACACCAACCC ACTATCTGCT CAGCAGTTGC CCGGGGCTAC CAGGCTGTGT TTACTTCCGT 480  
 TGCTAGGCTT TGTGTGAGGA TGAGCAGCA CCGGATAGAC AGGAAAGGAG TCTGTGCTAA 540  
 CCGGTATGCC GTGTTCACGG GTGTCAACCA GTGATATCGG TTGTGATGAG CCGGAGTCT 600  
 GCGCCGGGAC GGGGCGCAGC TGTGTATCAG CAGCGGGAAG CAGCAAGAAC TGGACCGGCG 660  
 70 CATGGCCAAG CTCAGGGGGG AGGGGGCTGAG TGTGGCGGCG ATTGTGTGCC AOSTGGGGAA 720  
 GGCTGAGGAC CCGGAGCAGC TGGTGCCCAA GGCCCTGGAG CACTGTGGGG GGTGTGACTT 780  
 CTCTGTGTGC AGCGCAGGGT TCAAGCCTCT GTTAGGAGAC ACTCTGGGGA CCAATGAGCA 840  
 GACTTGGGAC AAGATGAGG GTGTGAAGCT GAATGTACCA GCGCTCTGCG TGAGCAGATT 900  
 GCTGCGCTAC ATGAGAGACA GGAAGGGTGC TGTCTACTCT GTCTCTTCCA TTGACAGCTA 960  
 75 TAATCCAGTA GTGGCGCTGG GTGTCTACAA TGTCAACCAAG ACAGCGCTGC TGGGTCTTCA 1020  
 TAAAGCACTG GCATTGGAGC TGGCCCCCAA GGACATCCGG GTAAACTGCG TGGTTCAGG 1080  
 AGATTATAAA ACTGACTTCA GCAAAGTGT TCAATGGAAAT GAGTCTCTCT GGAAGAACTT 1140  
 CAGGAACACT CATCAGCTGC AGAGGATGGG GAGATGCAGG GACTGTGCAG GAATGTGTTT 1200

CTCTCTGTGC TCTCCAGATG CCAGCTACGT CAACGGGGAG AACATTCCGG TGGCAGGCTA 1260  
 CTCACCTGGG CTCTGAGAGG AGTGGGGGGG GCTGCGTAGT TGTGTGCCCA GCCCAGGAGC 1320  
 CTGAGGGGGT GTCTAGGTGA TCAATTGGAT CTGAGCAGCA GTCTGCATT CTGCCAGACT 1380  
 AGCAATTGG GGGCTTACT A TGCTAGGCT TGAAGAAAGA GAAAAACGCT TCGCAATTCT 1440  
 CC

Seq ID NO: 76 Protein sequence  
 Protein Accession #: NP\_005785

1 11 21 31 41 51  
 | | | | |  
 10 MLSSAVARQV GWFHPCARLS VRMSSTGIDR KGVLANRVAV VTGSTSGIGF AIARRLRDQ 60  
 11 AHVSSSRKQ QNVDMARAKL QGEGLSVAGI VCHVGKAEDR BQLVALEH CGGVDFLVC 120  
 15 AGVNLVGSST LGTSEQWDK ILSVNVKSPA LLLSLLPYM ENRRGA VILV SSIAAYNPVY 180  
 16 ALGVNVSKT ALLGLTRILA LELAFKDIRV NCVVPGIUK DFSKVFHNE SLWKNFKEHH 240  
 17 QLQHGSESD CAGHVSFLCS PDASYVNGEN IAVAGYSTRL

Seq ID NO: 77 DNA sequence  
 Nucleotide Accession #: NM\_002425  
 Coding sequence: 26-1453 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 | | | | |  
 25 AAAGAAGGTA AGGGCAGTGA GAATGATGCA TCTGTGCTCT TGTGTCTGCC 60  
 26 AGTCTGCTCT GCTATCTCT TGAATGGGGC AGCAAAAGAG GAGGAGTCCA ACAAGGATCT 120  
 27 TCCAGCAAAA TACTATGAAA AGTACTACAA CCGGAAAAGG GATGTGAACG AGTTTAAAG 180  
 28 AAAGCAAGT AATCTAGTGT TTAATAAAAT CCAAGGAATC CAGAAAGTTC TTGKGTGGA 240  
 29 GTGACAGGG AGCTAGAGCA CTGACACICT GGAGGTGATG CGCAACGCCA GGTGTGGAGT 300  
 30 TCGTGACGTT GGTCACTICA GCTCCTTCTT TGGCATGCCG AAGTGGAGGA AAACCCACCT 360  
 31 TACATAACGG ATTGTGAATT ATACACCAGA TTTGCCAAGA GATGCTGTG ATTCTGCCAT 420  
 32 TGAGAAAGCT CTGAAAGTCT GGGAGAGAGT GACTCCACTC ACATTCTCCA GGCCTGATGA 480  
 33 AGGAGAGGCT GATATAATGA TCTCTTCCG AGTTAAAGAA CATGGAGACT TTAATCTTT 540  
 34 TGATGGCCA GGCACACGTT TGGCTCATGC CTACTACCT GGAACGGGGC TTATGGAGA 600  
 35 TATTCACITT GATGATGATG AAAAATGGAC AGAAGATGCA TCAGGCACCA ATTTATCTCT 660  
 36 CGTGTCTGCT CATGAACITG GCCACTTCTT GGGGCTCITT CACTCAGCCA ACACCTGAAGC 720  
 37 TTGATGTAC CCACCTTACA ACTCATTAC AGAGCTGGCC CAGTTCGCGC TTTCGCAAGA 780  
 38 TGATGTGAAT GGCATTCAGI CTCTCTAGG ACCTCCCTCT GCTCTCATG AGGAACCTCT 840  
 39 GTGCCACCA AATCTGTCT CTTCGGGATC TGAGATGCCA GCCAAGTGTG ATCTCGCTTT 900  
 40 GCTCTCGAT GCCATCAGCA CTCTGAGGG AGAATATCTG TCTTTAAAG ACAGATATT 960  
 41 TTGGGCAAGA TCCCATGGA ACCCTGAACC TGAATTCAT TTGATTCTG CATTTTGGCC 1020  
 42 CTCTCTCCA TCATATTGAT ATGCTGCATA TGAAGTTAAC AGCAGGGACA CGTTTITAT 1080  
 43 TTTTAAAGGA AATGAGTCTT GGGCCATCAG AGGAATGAG GTACAAGAGC GTTATCCAG 1140  
 44 AGGCTACAT ACCCTGGTGT TTCTCCAC ACATAAGAAA ATTGATGAGC GTTGTCTGA 1200  
 45 CAAGGAAGAG AAGAAAACAT ACTTCTTTC AGCGACAAA TACTGGAGAT TTGATGAAA 1260  
 46 TAGCCATGCC ATGGAGCAAG GCCTCCCTAG ACTAATAGCT GATGACTTTC CAGGAGTTGA 1320  
 47 GCCTAAGGTT GATGCTGTAT TACAGGCATT TGGATTTTTC TACTCTTCA GTGGATCATC 1380  
 48 ACAATTGAG TTGACCCCA ATGCCAGGAT GGTGACACAC ATATTAAAGA GTACACAGCT 1440  
 49 GTTACTTGC TAGGCGGAGT AGGGGGAAGA CAGATATGGG TGTTTTAAAT AAATCTATA 1500  
 50 ATATTATCT TTTTSSSTVNP EFHILIBAT WPLFLPYLA AYVNSRDTY FFEKGNFPA 1560  
 51 GAAGAAGATG AGCCTTCCAG ATATCTGCAT GTGTCAITGA GAAITGTCTT GGAATCTTCT 1620  
 52 ACTTGCTTTT GAATTGCAC TGAACAGAAT AAGAATACT CATGTGCAAT AGGTGAGAGA 1680  
 53 ATGTATTTTC ATAGATGTGT TATTACTTCC TCAATAAAAA GTTTATTAT TGGGCTGTTC 1740  
 54 CTT

Seq ID NO: 78 Protein sequence  
 Protein Accession #: NP\_002416

1 11 21 31 41 51  
 | | | | |  
 60 MHLAFVLLC LPVCSAYPLS GAAKEEDSNK DLAQVLYEK YNLEKDVKQF RKKJSDNLVK 60  
 61 KIQGMQKFLG LEVIGKLDLD TLEVMRKPRC GVPDVGHPFS PPGMPKWRKT HLTYYRINY 120  
 62 PDLPRDAVDS AIEKALKVWE BVTPLTFSRL YEGEADDMIS FAVKHBODFY SFDGPHSLA 180  
 63 HAYFPQGLY GDHDFDDEK WTEDASGTL FLVAAHELGH SLGLFHSANT EALMYPLYNB 240  
 64 FTLEAQRLS QDDVNGHQL YGFFPASTBE PLVPTKSVPS GSEMPAKCDP ALSFDAISTL 300  
 65 RGEVLFKEDR TTFIRSRVNP EFHILIBAT WPLFLPYLA AYVNSRDTY FFEKGNFPA 360  
 66 IRNIEVQAGY PRGHITLGFY TPIRKIDAAV SDKEKKKTYF FAADKYWRFD INSQSEMQGF 420  
 67 PRIADDFGY YEFKVDVAVL AGFFVYFSP SSQFEPDFNA RMVITBLKSN SWLHC

Seq ID NO: 79 DNA sequence  
 Nucleotide Accession #: NM\_000493.1  
 Coding sequence: 1-2043 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 | | | | |  
 75 ATCTGCGCAC AAATACCTT TTTGCTGCTA GTATCCTTGA ACTTGCTTCA TGGAGTGTT 60  
 76 TACCTGGAAC GATACCAACG GCCACAGGCG ATAAAGGCC CACTACCCA CACCAAGACA 120

CAGTCTCTTCA TTCCCTACAC CATAAAGAGT AAAGGTATAG CAGTAAGAG AGAGAAAGT 189  
 ACTCCGCTGTC CAGTCTCTTCA CAGGAGCTC CAGTCTCTTCT TGGACACCA 248  
 GGAAAAACCA GCTACGGAAO TCTTGAGCTC CAAGGAGAGC CAGGGTTGCC AGGACCAACG 300  
 GGACCATCA9 CTTGATGGGAA ACCAGGTTGTG CCAGGACTCC CAGGAAAAAC AGGAGAGACA 320  
 5 GGACCATATG GACCAAAAGG AGATGTTGGA CCAGCTGGCC TACCAGGACC CCGGGGCCCA 460  
 CCAGGACCAC CTGGAATCCC TGGACCGGCT GGAATTTCTG TGGCAGGAAA ACCTTGACAA 480  
 CAGGAGCCCA CAGGAGCCCA AGGACCCAGG GUCTTTCCTG GAGAAAAGGG TGCACAGGA 540  
 GTCCCTGGTA TACAGAGCCA GAAAGGGGA AATGGATATG GTGCTCTGTG TCGTCACAGT 600  
 GAGAGGGCTC TTCCAGGGCC TCAGGGTCCC ACAGGACCAT CTGCGCCTCC TGGAGTGGGA 620  
 10 AAAAGAGGTT AAAATGGGTT TCCAGGACAG CCAGGCATCA AAGGTGATAG AGGTTTTCGG 760  
 GGAGAAATGG GAACCAATTGG CCCACCAGCT CCCCAAGGCC CTCTCTGGGA ACAGAGGGCA 780  
 GAAGGCATTG GAAGCCAGG AGCTCTCTGA GCCCAGGCGC AGCCAGGAT TCCAGGAACA 840  
 AAAGGCTCTC TCGGCTCTC AGGATAGCTT GGGGCCCA G GUCTCTGTG CTTTGGGAA 900  
 CCAGGCTTCC CAGGCTGTA GGGGAAAAGA GAACTGCTG GCTTCTGTG GGGTCCAGT 960  
 15 GCCAAAGGGG AACCAAGGCC ACAGGTTCTT CTGGGGAAGC CAGGTTCTGAC TGGACCCCT 1020  
 GGGAAATATG GACCCCAAGG ACCAAAAGGC ATCCCGGTTA GGCATGTTCT CCCAGGCCCT 1080  
 AAAGGTGAGA CAGGCGCAGC TGGGCTCTCA GATACCTCTG GGGCTAAGGG TGAAGGGTT 1140  
 TCCGCTGGGT CAGATGGA AACACAGGTTCC CAGGAAAC CAGGTTGGA TGGTCTAAG 1200  
 GTGAACCCAG GTTACCAAGG TCCAAAAGGT GATCTGGAG TGGAGGACC TCGTGTCTC 1260  
 20 CCAGGCGCTG TGGGCCAGC AGGAGCAAG GGAATGCCG GACACAATGG AGAGGCTGGC 1320  
 CCAAGAGGTT CCGCTGGAA TACAGGTAAT AGAGGCCCTA TGGGCCAAC AGGCATTCCA 1380  
 GATGTTCCCT GTCTCAAGG GGTACCAAGA AGTCCGGTCT CTCTCGGCC AGCTGTCATA 1440  
 GCAACTAAGG CCGCTCAATGG ACCCAACGGG CCACAGGGCC CTCACAGTCC TAAAGGGCCC 1500  
 TGTGGAGAGT GTGCTCTCTC AGGCGGCTG AGGCGCTTCA GGCACAGG TCAAGGCTCT 1560  
 25 ATGGCTGAGG GTTTTATAAA GGCAGGCCAA AGGCCAGTC TTCTGGGAC CCGTCTTGT 1620  
 AGTGCCAACC AGGGGGTAAC AGGAATGCC TGTGTGCTT TTACTGTTAT TCTCTCCAAA 1680  
 GCTTACCCAG CAATAGGAAC TCCCATACCA TTGTATAAAA TTTTGTATA CAGGCAACAG 1740  
 CATTTATGAC CAAGGACTGG AATCTTTAT TTGCAGATAC CAGGAATATA CTATTTTCA 1800  
 TACACCTGAC AAGTAAATGAG GACTCTATTT TGGTAGGCC TGTATAAGA TGGCACCTC 1860  
 30 GTAAATGACA CTTATGATGA ATACACCAAA GCTCATCTGG ATCAGGCTC AGGAGTCCC 1920  
 ATCATGATC TCACAGAAA TACCAAGGTT TGGCTCCAGC TTCCCAATGC CGAGTCAAA 1980  
 GGCTATACT CCTCTGAGTA TGTCCACTCC TCTTTCTCAG GATTCTAGT GGTCTCAAT 2040  
 TGAATACACC CCACAGAGCT AATCTAAAT TTGTGCTAGA AAAAGGATTC TCTAAGCTCA 2100  
 35 CCCCACCTTA CAAATGTCAT ATGGAGGTAG GCTGAAAAGA AGTAAATT TATTCTGTA 2160  
 AATACAGATT TGACCTATCA GACCAACAAA CTTCTCCCTT GAAAGTGAC CAGCAAGCTA 2220  
 AAAACGTTAT TGAAGCCTCT CTGGAATTTC TAGTTGCAA TCTTAAGGCT CTTAAAGTT 2280  
 TCTTCAATA TAAAAAATA TCACCAAGA AGTCTGCTA TGTATAAA ACACAACAAA 2340  
 AACAAAGACA ACACAAAAAA AATATAAAAA AAAAAACAGA AATAGAGCTC TAACTTATG 2400  
 40 GAAATTTGAT TGGAGAAACT CCGCATTTCC TTTTAAAAAA AGCCTGTTTC TAACATAAA 2460  
 TATGAGAACT TCTAGAGAAC CACAGGAG TATCATATA CTTTGTAGAA CTTAATAAT 2520  
 TGAATATCA AATTAAAAAG ACACGTATC CCTTAAAAA TTTCTGATGG TGCATCTAC 2580  
 TGAAGGCTGT ATGGGCCCTT TCATCAATAT CTATTCAA ATACAGGTGC ATATATACIT 2640  
 GTTAAAGCTC TTATATAAAA AAGCCCCAAA ATATTGAAAT TCATGTGAA TGAAGGTTG 2700  
 45 TTTCATCAAT GAACTTTTC AAAAGCTTTT TATGATTGCA GAGAAGCTTT TTATATAACC 2760  
 AGCATACCTT GGAAACAGGT ATCTGACCTA TTCTATTATA GTTAAACAAA GTTGATTTAA 2820  
 TTTGATTTGT TTATCTCTT ATGATCTT ATGTGATAT ATTTCTGGA TTACAGAAC 2880  
 ATTAGCAGAT GTACCTTTGT CTTCCCATC AAGTGAAGT ATAAATTACA CTGAGGRTT 2940  
 CAAAATTCGA CTAGAAGTGG AGATATATTA TTTATTATG CACTGTACTG TATTTTATA 3000  
 50 TGTCTGTTA AAATCTTTAA GCTGTGCCCT ACTTATTAAC GCACAAAAATG TTTACCTAC 3060  
 TCCTTATTA GCACACATA AAATAACATC AATAGATTTT TAGGCTGAAT TAAATTGAA 3120  
 GCACAGATT GCTCTATGCA ACCATCTTT CAAGGCTTTT CATTGCACAC AATATGAAA 3180  
 CATCAATAGA TTTTATGG

Seq ID NO: 80 Protein sequence:

Protein Accession #: NP\_000484.1

1 11 21 31 41 51  
 60 MLPQIPFLLL VSLNLVHGV YARYQYPTG IKGPLPNTK QFPIPYTKS KGIAVRGEQ 60  
 TPQHPGAPR RGHPPGSPFP GKPGYQSKL QSEKPLPFP GSAVGRKGV PGLPKKPR 120  
 GPYQKGVDP FAGLPKPRP GAGPQKPCA GSVQPKPQF QKPTGAPRPF GPPKRGAPG 180  
 VPMNGNQKRG MGYGAPGRPG ERGLPQPGQ TPGSPQVPG KRGNVVPVQF GKUKDRGPF 240  
 65 VPMGPIGPPF PQQPPGPRPF BGIGKPGAAG APQQPGIPGT KQLPGAPGIA GPPPGPGFK 300  
 PGLPLKGER GFAQLPGGPG AKGEQQA GL PKPGLTQFP GNMGPQPKRG IPGSHLPQF 360  
 KQETGPAQFA GYPAKGERGU SPGSQKPFYF PGKPLGDPK GNPLPGPKG DPVQVQPL 420  
 PIVYVPAKAK GMPGVNGIAG PNGAPRPFIT NGKNGPFI GPFGSKDQPG SPQPPFAGI 480  
 ATKQLNGPT PPQPGPRPF SGEPIPLFPF GPFGPGQAV MPFGFKRAGQ RFLSGLPLV 540  
 70 SANQVYTGMP VSAFTVLSK AYPALGTFP FDKILYNRQY HYDPRTOIFT CQFGLIYFS 600  
 YHVHVKGTHV WVGLYKNGPT VMATYDEYTK GYLDQASGA IDLTENDQV WLQLPNAESN 660  
 GLYSSEYVHS SPSFGLVAPM

Seq ID NO: 81 DNA sequence

Nucleic Acid Accession #: NM\_000786

Coding sequence: 332-1861 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51

CGCGATTCTC AGGGATTGAT CCGGCTCTTC AGGTAAGTTA TCTTCGGGCC CCGTACCACT 60  
 GTGGCCACAGG CGCAGGCCGCC TTCTCTAGGT GCCATCTCCC GCGCAGAAGA CCACGGCTTC 120  
 ACAAGTGTTT ATTTAAGGGC GTGGCCACGG GAACATCCCG CCCATTCTGT TGACGCACGG 180  
 5 GTGTGGCGGC GCGACACCCG AGGGGTGGGG CTGGGTTTAG TAGGAGACCTT GGGGCAAGGC 240  
 CCGCTGTGGA GCGACATCTG CCAGCTCTCT TGGTTCGCTT GATTGGGAGG AGGGGTGGGG 300  
 ACCTCGGCTT TCACTGTTTC GCACGGAGTG AATGGCGGGC GCGGCTGGGA TGCTCTGCTC 360  
 GGGCTTGCTG CAGCGGGGTG GGTGGGTGCT GGGCCAGGCG ATGGACAAGG TGACACGGCG 420  
 CAACCTCTTG TCCATGCTGC TGATGGCGCT GCGCTCTACC CTCAGGCTGG TCTACCTGAT 480  
 10 CGGTCTGGCC ATGATGCG CTGGGCGATCG CATGACATT GGGAAAGTGC GAATGAGATT 540  
 TTCTGCGGA ATGATGCG CTGGGCGATCG CATGACATT GGGAAAGTGC GAATGAGATT 600  
 TCTGAAAAAT GCATATGAGA AGTATGGAAC TGATTTAGT TTACCATATG TAGGCAAGAC 660  
 ACTATCTTAC CTCTGGGGGA GTGATGCTGC TGCACTGCTT TTTAATAAGT AAAATGAAGA 720  
 CTGGAATGCA GAAGATGTCT ACAGTCCGCT GACACACCTT GTGTTGGGA AGGGAGTTGC 780  
 15 ATACGATGTG CCTAATOCAG TTTCTTGGG GCGAGAAGAAA ATGTTAAAAA GTGGCTCTAA 840  
 CATAGCCGAC TTAAACAGCG ATGTTTCTAT AATTGAAAAA GAAACAAGG AATATCTTGA 900  
 GAGTGGGGA GAAAGTGGAG AAAAANAATG TTGTGAAGCT CTCTTGAAGC TCATAATTG 960  
 AACACGTAGC CATTGTTTGC ATGGAAAGGA AATCAGAAGT CAACCTAATG AAAAGGTAGC 1020  
 GACAGCTGAT CGAGATTGGG ATGGAGGTTT CAGCCATGCA GCGTGGCTCT TACCAGGTTG 1080  
 20 GCTGCTTTG CTCAGTTTCA GACGCAAGGA CAGAGCTCAT CGGGAATCA AGGATATTTT 1140  
 CTATAAGGCA ATCCAGAAGC GCAGACAGTG TCAAGAAAAA ATTGATGACA TTCTCEAAAC 1200  
 TTACTAGAT GCTCACTACA AGGATGGGCG TCCCTTGACT GATGATGAAG TAGCGAGGAT 1260  
 CTTATAGGA TTACTCTTGG CAGGGCAGCA TACATCTCTA ACTACTAGTG CTTGATGGG 1320  
 25 CTCTTTTGG GCGCAGGACA AAACACTTCA AAAAAAATGT TATTAGAAC AGAAACAGAT 1380  
 CTGGGAGAG AATCTGCTCT CTTTAACTTA TGACCAGCTC AAGGATCTAA ATTACTTGA 1440  
 TGGGTGTATA AAAGAATCAT TAAGCATTAG ACCTCTCTAT AAGTATGTA TGAGAAATGG 1500  
 CAGACCTTCT CAGCATGTGG ATGAGTATAC CATCTCTCA GCGACATGCG TTGTGTTCT 1560  
 TGGCACTGTC AATCAAGGAC TTAAGAGCTC ATGGGTGATA GCGCTGCACT TTAATCTGGA 1620  
 30 TGCTCACTTA CAGGATAACC CAGCATCAGG GGAAAGAATT GCCTATGTGC CATTGGAAGC 1680  
 TGGGGTCAT CGTTGATTT GGGGAAATTT TGCTATGTT CAATTAAGA CAATTGTGCT 1740  
 CACTATGCTT CGTTTATA TGATTTAGTCT CATTGATGGA TACTTCCCA CTGTGAATTA 1800  
 TACAACATGT ATTCACACCC CTGAGAACC AGTTATCGGT TACAACAGGA GATCAAAAGT 1860  
 35 AAAAGGTTG CAGGAGAGGA AATATGTGTA TTAATCTGTA TTAATCTGTA GGCATCTGAA 1920  
 GAGATGAAG TGATCAAAAC AACTCTTGTA GTTACTGTT TTTTAAAGTG TGAATCTTA 1980  
 AAAGCCAGTT TATGATTAG GATTTTGTTA ACTGAATGTT TCTATCAAT ATAATAGCAT 2040  
 TCCACATCT TCTAATAGT TATGATACT ATACATGTGC TTACGGAAG TTCTTTGGT 2100  
 AAACAATGT TGAGGGGGGA TCTAGTGTA TGGCAGATTC TAAATAATAT AATTTCGAGA 2160  
 40 TAGTAAATTT AAGGATATCT AGGCTCTTGG CCAATAAATG TCAGGGATAT CAATCTTGG 2220  
 ACTAGCTCGT CAGAGTATAA AGAATAAAAA TCCCATGTAG ATACTTGGAA ACCAGAGTT 2280  
 ATATTATTT ATGAGGCGCA TTATGTGTTG TGTGAGGATG GAAGGTAGG GAATAATGGA 2340  
 ACATCTAAAG CTTTGAATAA GAGAATACTA ATGTGTTTGG TATGATGATA CTCAGAAATG 2400  
 GAGATATTT AGGAAAAAGA AATCCTTTGG AATTTTAACT AAAATCACTG CATATGGGAA 2460  
 ATTAAGAGAT CCAGAGCACT ATTTGATTAAG AGTCTCTAAA AATAATGTAA TTATTAATGC 2520  
 45 TAAAGACTGC TCAATGATCT TGATCTAATT ACTAAATAAA TTACATATTT ATTACTGGA 2580  
 TAAATATGTA TCACTGTCTA CAAGGTGACA TTATGTGGA AGTCCAAAGT CAAGTCTCTA 2640  
 GGGGATAATT TTGTTTGGG CTCAGTTGTT CCGCTGCTCC TTTTITTTT TTTTITTTT 2700  
 TTGAGATGGA GTCTGCTCT GTTGGCCAGG CTGGAGTGGA GTGGTGGAT CTGAGCTCAC 2760  
 TGCACTCTCT GCGTCCGGGG TTCAAGCAAT TCTCTGCTC AGCTCTCAA GTAGTGTGGA 2820  
 50 TTACAGGCAC GCGTCCAGCAT GCGTGGTATA TTTTGTGTT TTTAGTGAAG TCGATGCT 2880  
 TCACTATTT GCGTCAAGTG GTCTTGAATC CCGTAGGCTC GTAGTCCAC CCGGCTGGC 2940  
 CTCCCAAAGT GCTGGGATTA CAGGCATGAG CCACGCCACC TTGCGCTCCC TGCTTCTCT 3000  
 CTAGAAATCA ATTAGGGATG TTTGTTACTA CTCATATGTA TTAACAACAGT TAAACAATCT 3060  
 TTTCTTTT TAAATGTGAG ATCAGTGAAC TCTGTTTITA AGATAATGCT AAACAAGGCT 3120  
 55 CTGAGATTA ATAAAGTCTG TCAAGCATCT TGTGTTGAG AATCACTATA TCTGATCTA 3180  
 TCTAATG TAACTCCGGG CCTAATCTTA TGAGATGGCT GAAAAAATAC CACATCTGCT 3240  
 GTTTTCACTA GGTGATGCA AAATAATTTG CTTATGTAT ATTACAGTCT TTTTAAAC 3300  
 ACTGGAAGAC TCAATGTAAA CTCTAATTGT GAAGGCAGAA TCTTGCTAAA TTTTICAGAT 3360  
 TAAATTTCT TTTGAAAAAA T

Seq ID NO: 82 Protein sequence:  
 Protein Accession #: NP\_000777

1 11 21 31 41 51  
 65 MAAAAGMLLL QLLQAGSGVL GQAMEKVTGG NLLSMLLIAC AFTLSLVYLI RLAAAGHYQL 60  
 PAGVKSPPYI FSUPLGLHA IFGKSPFER LENAKEYKG VFSEFMVGKT FTYLLGSDAA 120  
 ALLENSKNEI LNAEDVYSRL TTPVFGKGA YDVPNPVLE QKKMLKSLN IAHFKQHVSI 180  
 IBKETKEYFE SWGESGEKNV FEALSELIL TASHLHGKE IRSQLEKVA QLYADLDGOF 240  
 70 SHAAWLLGW LPLPSFRRRD RAHREIKDF YKAIQKRQS QEKIDILQL LLDATYKDR 300  
 PLTDEYAGM LELLLAGQH TCTGTBAWNG FLAARQKTLQ KCKYLSQTY CGBLPLTGY 360  
 DQKDLMLD NCKEISLEL RFIMIMMMMA RTKQTVAGYI TPGHGVYCS PTYNGRIKDS 420  
 WYKELDNPD RYLQDNPSG EKFAVYVFGA GRIRCGENF AYVQCITWS TMLRIYEPDL 480  
 IDGYFTPNY TMTHTPENP VIRYKRKSK

Seq ID NO: 83 DNA sequence  
 Nucleic Acid Accession #: NM\_006512  
 Coding sequence: 64-336 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 5 AATTC TAGAA GTCCAAATCA CTCATTGTT GTGAAAGCTG AGCTCACAGC AAAACAAGCC 60  
 ACCAATGAAGC TGTGCTGTG TCTCTGTGTG GTACAGCTGG CCTCTGTGTG CTACCAAGGCC 120  
 AATGCCGAGT TCTGCCAGC TCTTGTCTT GAAGCTGTAG ACTTCTCTCT CATATGTGAA 180  
 CCTCTGTCTA AGTTAAGTCT TGCAAAATTI GATGCCCTTC CGGAAGCTGT TGCAGCCAA 240  
 TTAGGAGTGA AGAGATGCAC GGATCAGATG TCCCTTCAGA AACGAAGCTT CATTCGGGAA 300  
 10 GCTCTGTGTA AAATATTGAA GAAATGATGT GTGTGCATG TAAAAACTTT CATCTGTGTT 360  
 TCCACTGTCT TCAATAGACA CCTGATCTT CACTGCAGAA GTTAAAGGTT TCAACGTCTT 420  
 GCTTAAATA ATCACTTGTCT CTAC

Seq ID NO: 84 Protein sequence  
 Protein Accession #: NP\_00542.1

1 11 21 31 41 51  
 15 MKLSVCLLLV TLALCQYQAN AEPFALVSE LLDFFPISFP LFKLSLAKFD APPEVAAKL 60  
 20 GVKRCDQMS LQKRSLIAEV LVKILKKSV

Seq ID NO: 85 DNA sequence  
 Nucleic Acid Accession #: NM\_002317.1  
 Coding sequence: 231-1484 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 25 GGGCCAGGAC TGAGAAAGGG GAAAGGGAAG GGTGCCACGT CCGAGCAGCC GCCTTGACTG 60  
 GGGAAAGGTC TGAATCCAC CCTTGCAATT GCTTGTTGGA GACTGAGATA CCGCTGCTCC 120  
 GCTGCTCTTC TTGTTGTGAG ATTCTCCCTT CCTCAGCTGG ATTGAGCCGC GTTTTATT 180  
 30 TTCTGTGAGC CAGCTCTCTC TCGAGCGGGG TCAACTTGGC AAAAGGAGTG ATTCGCTTGG 240  
 CCTGAGCCGT GCTCTGCTCT GGGCCTTTGC AGCTCTGGCG GCTAGTGAC CTGCCGCCCTC 300  
 CGCGCGCGGG CCAACAGCAG CCGCCCGCGG AGCGCCCGGC GCTCTCGGCG GCTCTGGCGC 360  
 AGCAGATCCA ATGGGAGAAC AACGGGACGG TGTTCAGCTT GCTGAGCCTG GCTTCACAGT 420  
 ACCAGGCTCA GCGCCCGCGG GAACCGGCGG CGGCGCTCCG TGTGTGACCC AAGGCTCCCG 480  
 35 CCGAGCAGCC CGGCACTCCG ATCTCTGTGA TCCGCGACAA CGGCAACCGC GCGGGCGGAA 540  
 CGCGGACGCG CGGCTCATCT GGAATCAGCG CTGTCGCGCC CAGGCCACCC GCGGCTCACT 600  
 GTTCCAAAGC TGGCTACTCG ACATCTAGAG CCGCGAAGC TGGGCGCTCG CGCGGGGAGA 660  
 ACCAGACAGC CCGCGGAGAA GTTCTTGCTC TCAATAACTT CGGCGCGGCC AGCCCGCTGG 720  
 AGCGCATGCT GCGCGACGAC CTTCAACAC CTTACAAGTA CTCTGAGCAC AACCTTATT 780  
 40 ACAACTACTA CGATCTACTT GAAAGGCCCA GACCTGGGGC CAGGTACCG CCGCGATACG 840  
 G CACTTGCTA CTTCACGTAC GGTCTCCGAG ACCTGTGGCG CGAGCCCTAC TACATCCAGG 900  
 CGTCCACGTA CGTGCAGAA ATGTCCATGT ACAACCTGAG ATGCGCGGCG GAGGAAAACT 960  
 GTCTGGCCAG TACAGCATAC AGGCGAGATG TCAGAGATTA TGATCACAGG GTGCTGCTCA 1020  
 GATTTCOCA AAGAGTGAAA AACCAAGGGA CATCAGATT CTACCCAG CGACCAAGAT 1080  
 45 ATTCTGGA ATGGCAGAGT TGTCTCAC ATTACCAGAG TATGAGTAG TTATGCCACT 1140  
 TTGACTGTCT TGAATGCAAC ACCGAGGAGA GATGGGCTGA AGGCCACAAA GCAAGTTTCT 1200  
 CTTTGGAAGA CACATCTGT GACTATGGCT ACCACAGCGC ATTTGATGT ACTGCACACA 1260  
 CACAGGAGAT GAGTCTCTGC TGTTATGATA CCTATGGTGC AGACATAGAC TGCCATGTGA 1320  
 50 TTGATATTAC AGATGTAAAA CTGGAAGACT ATATCTTAAA GGTCAAGTGA AACCCGAGCT 1380  
 ACCTTGTTC TACAACTGAC ATATGTGCGA ATGTTGTGCG TGTGTGACCT GCTTACAGAC 1440  
 GACATCATTC GTATGCTCA GGTGTGCACA TTTCACCGTA TTGAAAGGCA AAGCAAAACT 1500  
 CCCAATGGAT AAATCAGTGC CTGTGTGTCT GAAGTGTGAA AAAATAGACT AACTTCAGTA 1560  
 GGATTATGT ATTTTGAAAA AGAGAACAGA AAACAACAAA AGAATTTTGT TTGAGACTGT 1620  
 55 TTTCATAAC AAAGCACATA ACTGAGATT GAACGCTTAA GTCATCTA CTGGGAAAT 1680  
 TTTTAAATGT TATTATTAC ATCAGTTGT GAATTAACAC AGTGTTCAC TCTGTGAAAT 1740  
 ACATATTGA CTCCTTCAA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 86 Protein sequence  
 Protein Accession #: NP\_002308.1

1 11 21 31 41 51  
 60 MRFAWTVLLI GPLQICALVH CAPPAAGQQQ PFREPPAAPG AWRQQIQWEN NGQVFSLSL 60  
 65 GSQYQQRRR DPGAAPGAAA NASAQFPTP ILIRIDNRTA AGRTRTAGSS GVTAGRPPT 120  
 ARHWQAGYYS TSSAREAGPS RAENQTAFGE VPALSNLRPF SRVDGMVGD PYNPKYSYSD 180  
 NPTFNYDTY ERBQGRGVR PGYGTQYRY GLDFVADYF VTQASTYVQK MSMYNLRCAA 240  
 HENCLASTAY RADYRDYDR VLLRFRGVR NQTSIDLIS RPRYSWEWHIS CHQHYHMDSE 300  
 FSHLYLLDAN TORRWBGHK ASFLCEDTSC DYGYHRRFAC TAHTQGLSPG CYDTYGAD 360  
 QWIDITDVK PGNYILKVSF NPSYLVPSDF YTNVYVRCDI RYTHHAYAS GCTISPY

Seq ID NO: 87 DNA sequence  
 Nucleic Acid Accession #: NM\_006419.1  
 Coding sequence: 91-420 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 75 TTGCGCACTT GGGAGAAAGT GTTTGAAAAA ACTGACTCTG CTAATGAGCC TGGACTCAGA 60

GCTCAAGTGT GAACCTTACC TCCAGACAGA ATGAAGTTCA TCTGCACATC TCTGCTCTC 120  
 ATGCTGCTGC TCAACACCCCT CTCTCCAGTC GAAGTGTTTC TGGAGGTCTA TTACAACAAGC 180  
 TTAGAGTGTA GATGTGTCCA AGAGAGCTCA GTCTTTATCC CTAGACGCTT CATGTATCGA 240  
 ATTCAAAATCT TGCCCGGTGG GAATGTTGT CCAAGAAAAAG AAATCATAGT CTGGAAGAAG 300  
 AACAAGTCAA TTGTGTGTGT GGACCTTCAA GCTGAATGGA TACAAGAAT GATGGAAGTA 360  
 TTGAGAAAAA GAAGTTCTTC AACTCTACCA GTTCCAGTGT TTAGAGAAAA GATCCCTGTA 420  
 TGTCTGATTT TCCATCAAGA ACACCTGCAT TCTTCCCTTA TCCCTCTCT GGAATTTAGT 480  
 TTGTGCTTA GTTAAATCTT TTCCAGGGAG AAGAAGCTTC CCCATACAAA TAAAGCATGA 540  
 GAGCTATGTG AAAAAAACC TTGCAGGAGC TGAATGGGGCA AACTCAAGT TCTTCACCTA 600  
 CAGCACCCCTA TATACACTTG GAGTTTGCAT TCTATTTCAT CAGGAGGAA AGTTCITTTG 660  
 AAAAAAGTTA TCACTGTTA AGTAATACAG GATTATTTTG ATTATATACT TGTGTTTAA 720  
 TGTTTAAAT TCTTTAGAAA ACAATGTGAAT AGGAATTTAA GCTTCAATTT TGAACATGTG 780  
 GCTTGAATTA AGAAGAAAAAT TATGGCATAT ATTAAGAACGA GCCTCTCATG AAGAGACTCA 840  
 AAAGCTGCTT GGGAGGCGCA TGGAACTTGA GCCTGTCAAG AGGCAAGGA ATCCATGTAG 900  
 TAGATATCCT CTCGTTAAAA ACTCACTACG GAGGAGAAAT AAGTCTACT TTAAAGAAAT 960  
 TCTTTTATAA AATTTACTGT TAAAGATTAA TAGCATTCGA AGATCCOCAG ACTTCATAGA 1020  
 ATACTCAGGG AAAGCATTTA AAGGTGTGAT TACACATGTA TCTTTCACA CATTTCCTT 1080  
 GACAAACTTC TTACATCAAC ATCTTTTCA CTGACTTTT TTGTGGGGC GGGCGCGGG 1140  
 GGACTCTGTT ATCTAATTCT TTAATGATTC CTATAAATCT AATGACATTC AATAAAGTTG 1200  
 AGCAAACTT TTACTT

Seq ID NO: 88 Protein sequence:  
 Protein Accession #: NP\_006410.1

1 11 21 31 41 51  
 MKPISTILL MLLVSSLSPV QQVLEVVYTS LRCRCVQESS VFPRFIDR IQLFRNGNC 60  
 PKKEIIVWKR NKSIVCVDFQ ABWIKRMMEV LKRSSSTLP VFYFKRKF

Seq ID NO: 89 DNA sequence  
 Nucleic Acid Accession #: NM\_002652  
 Coding sequence: 37-477 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CTCTCTGGG ACACATGGCC TCTGTTTTC TCCAGCATGC GCTTGCTCA GTCUCTGTC 60  
 AGGGCCAGCC CTGCCACCGT GTCCTGTTT CTCTGCTGC AGTTGGGGGC CAACAAGCT 120  
 CAGGACACCA CTCGGAAGAT CATAATAAAG AATTTTGACA TTCCCAAGTC AGTACGTGCA 180  
 ATATGACGAAG TCACCTCAGT GCTTGAGATT CAACAGCAAT TGAAAGATG CAGTGTTGT 240  
 AAAAAGCTCA TCAATAGCG CATCCCTCTA CAAGGTGCAT TTAATCTATA GTATCTGCG 300  
 TGCTATGTTG ACBACAATCC AAAAACCCTC TACTGGGACT TTACACCCA CAGAAGCTGT 360  
 CAAATGCGAG CGGTGTTGA TGTTATTGCG GAATTAGCCA TCGGCCCTGA TGATGCTGCT 420  
 GTATCOCCA TCAAAAACAA CCGGTTTAT ACTATTGAAA TCTTAAAGT AGAATAGTGG 480  
 AAGCCCTGTC TGTGTGCAAC ACCCAATGTA TTCTCTTAA AGAAACTTGG CTGGAATTC 540  
 TCGTGTGTC TATAAATAA ACTCTTAAAC ATGCTT

Seq ID NO: 90 Protein sequence:  
 Protein Accession #: NP\_002643.1

1 11 21 31 41 51  
 MRLQLLFRA SPATLLVL LQLGANKAQD NTRKHUKNF DIPKSVRPND EVTAVLAVQT 60  
 ELKEBGMVVK YLSSIPLQG AFNYKYTACL CDDNPKTFYV DFYTNKTVQI AAVVDVIREL 120  
 GICPDAAVPI KHNRFYTI EILKY

Seq ID NO: 91 DNA sequence  
 Nucleic Acid Accession #: AK000341  
 Coding sequence: 85-973 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 GATAGCGCGG GCGACAGGGA CCGGCTAAC CTGGACAGCG CATCGCCGCG CGCCCGGGTC 60  
 GCGCGGCCAC AGCCGCTGCG GATCATGAAA CATCTAAAAG CTTTGTATGA TGAATCAAT 120  
 GCCTTTTGG ACAATATGTT TGGACCGGGA GATTCGAG TCAGAGGGTG GTTCACTGAT 180  
 GACTCTTACC TTCTACTCTT TTTCTTACT GTCATGTATC TGCCTCAAT ATGGCTGGGT 240  
 AACAGATATA TGAAGACAG AGCTGCTCTT TCTCTCAGG GTATCTGAC CTGTGTATAT 300  
 CTGGGAATCA CACTTCTC CGGTACATG CTGGCAGAGC TCACTCTC CACTTGGGA 360  
 GGAGCGTACA ACTTCAAGTG TCAAGATCTT ACCAGCGCAG GGAAGCTGA CATCGGGGTA 420  
 GCGAAGGTGC TTTGGTGGTA CTATTCTCC AAATCAGTA AGTTCCTGGA CACAATTTC 480  
 TGTGTTTTC GGAATAAACC GAGTCAGATT ACTTTCTTC ATGTATATCA TCATGCTTCT 540  
 ATGTTTAACT TCTGTGGTG TGTCTGAAAC TGGATACCT GTGGACAAG TCTCTTGG 600  
 CCAACACTGA AGCTTCTT CCAATCTCT ATGTACTCCT AGTATGGACT TCTGTGTT 660  
 CCATCTATGC ACAAGTATCT TTTGGTGAAG AAATATCTCA CACAGGCTCA GCTGGTGACG 720  
 TCTGTGCTCA CCAATCAGCA CCAATGAGC GCGTGTGTA AACCGTGTG CTTCOCCTC 780  
 GTGTTGCTCA TCTTCAGTC ATCTTATATG TCAACGTAG TCACTCTCTT CTTAAATTT 840  
 TATTGTTCAGA CACTCCGAAA AAAGCCCAAT AAGAAGAGTA TCGAAGAGC ACCGCGAGG 900  
 AAAGAAGTGA AGAATGGTTT TTCCAAAGCC TACTTCACT CAGCAAAATGG AGTATGAGAC 960

AAGAAGCAC AATAAAATG AGTAACAGAA AAGACACATA TACTAGCCTA ACAGATTGCG 1020  
 TTATTATAA GCALAACTG AATTGAAGTG TACATGTTTT AGGATAAAGT AATTTCTGTT 1080  
 GAGTTCAATA ATCAATTGTA CCGAGAATGT ATTAATATAT TGCTATTAGG TATCTGTTT 1140  
 AACTGAATGC TTGATCAGC ATTGAGGTGA TGCTCACTCT CGAGGACCTC AGAAGCTGGT 1200  
 CAGCTCTCT CTCCCTCCCT CCCACAGACT GAACCTTGC CAGAAAGCTG TCGTTATAAC 1260  
 GCCTTATAGC CATACACAGC CAGGAAGCTG GGAGCAATGT TTTCACAGA GAGTCTCAA 1320  
 ATAAAAAGG TTGTTGTCAG ATTAAATGT TTACACAAAA ATGTTATTA TATCTTAAT 1380  
 ACAGGTATG TTCTAACTTA TATTAAAGCA TAATGCAAGT GCATAATCA TCAATCTGTT 1440  
 OCTTTAGCAA TCAACCCGAC AAAATATATA AATGGGATCA TACACAAGAT ATAGAAAAAT 1500  
 CTAGCAAAAC TTCTCTTCT GTAAGCCAGA GTCTTGCTA TCAATTCOC ACAACCACT 1560  
 CTGATCTAA ATTTAGTGAT ATGGTAATGA AATGTGGATT TATTTAAAT ATAGTTATT 1620  
 CTAGGAGAAA AAAAATGCTT TCGCAAGAT TTCTAATTC ATGGAGCTG GATAGGATG 1680  
 TTCTCTGTT TCGTAATA TGATCTGTT CATGTGCC TCTTGCCA CTCAGCTAG 1740  
 GTTATACAGA TGCCATGCTC CACACCAGA CGAGGTATCA AATCTGGCTG CCGCTTACT 1800  
 TTCTGAGCAA GCACTGGAGT CACTCTGAC CTTTTCTT GAACATGCAT GCTGCTGAA 1860  
 TATGTATAA TCAGAACTAG CAGAAGTAG AGAGTGATGG GAGCAAAATA GGCCTAGAT 1920  
 TCGTCAACT TTTTGTGTA GCTCACTGT GAATATTAAC TCAGATCTCT GTTGTCACTC 1980  
 TTCAAGAGTT ATTTAAGTTC TGAAGCTGG GAGGAAAAG ATGAGTATGC TTGAAAAGT 2040  
 TCCAGCACTG AGCCGTGAAG CGGTCAATGAG CCAAGATAAA AAATGCCAT TTGGCAAACT 2100  
 CAGCACTCT GTTCCCTGCT CAGGTATATG OGATCTCTAC TGAGAAGCAA GCACAAAAGT 2160  
 TCCAGAAAGT ATTAATGAGT ATTTCTTTC TCCATAAGTG CAGGACTGTT ACTCACTACT 2220  
 AAACCTTACC AAGAATGGAA ACCAAGAATA TTCTCTGAAG ATTTTGTGAA AGATTAAAT 2280  
 ATACCTATA AATAAAACT TGTAGCTTC GATGAAGTCA AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 92 Protein sequence

Protein Accession #: BAA91096.1

1 11 21 31 41 51  
 MEHLKAFDE INAPLDMFG FRDSRVRGW FLDSYLPFF LTVMYLLSW LGNKYMKNP 60  
 ALSRLGILT LYNGLITLISA YMLAEILSLT WBGYNLQCG DLTSAGEADI RPAKVLVWVY 120  
 PXSVEPLDT IFVLRKKCT QITPLHYVHH ASMPNINWCV LNWPQGSQF POKPLNFSVH 180  
 ILMYSYGLS VFPMSHKYLV WKELYLTAQL VQVPLITHT MSAVVKPQF PFGCLGQSS 240  
 YMLTLVLFL NFYVQTYRKG PMKMDQBF AGKEVNGFS KAYFTANGV MNKJLQ

Seq ID NO: 93 DNA sequence

Nucleic Acid Accession #: NM\_000044

Coding sequence: 1115-3874 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CGAGATCCCG GGGAGGCCAC TTCTGGGAG AGCGGGACGG TCGGAGCAA GCCCACAGGC 60  
 AGAGGAGCGC ACAGAGGGAA AAGGUGCCGA GCTAGCCGCT CAGTGTCTGT ACAGGAGCGC 120  
 AAGGAGACCA CCAAGGCCAG CCCAGCCCGG CTCACGACAG AGCCAAAGCC TCTTCACGCG 180  
 CGCGGCTCT GAAGCCCGCT CCCGAGCTCG CCGTTCTCT TGGGTGAGG TTGTTAAAG 240  
 CTGCTAAGA CTCGAGGAAA GCAAGGAAA TGCTGGTAG GACTGACGCG TGCTTTGTC 300  
 CTCTCTCTCT CCACCCGCGC TCCCCCACC CTGCTCTCC CCGCTCCCG GTCTCTCTC 360  
 CCGCAGCTGC CTCATCGGC TACTCTCAG CAACCCCGCT CACCACCTT CTCGCCACCC 420  
 GCGCCCGCGC CCGCCTGCGC CCAAGCTCTG CAGCGCGAGT TGGCAGAGA GTAACCTCT 480  
 TTGGCTCGCA GCGGAGCGAC TACTGACATA TTGAAAGAAA GCGCTCTAGG AGCCAGCGGA 540  
 CTGGGAGCGC GCTTCAGCAC TCAGCAACAG ACCCGCTGTG TTAGAATTC GGGGAGAGA 600  
 ACCTCTGTT TTCCCGCACT CTCCTCCAC CTCCTCTGC CTTCCCAAC CCGAGTGGCG 660  
 AGCAGAGATC AAAAGATGAA AAGCGATCA GTCTTCAGT AGCCAAAAA CAAAACAAAC 720  
 AAAAAACAAA AAGCCGAAAT AAAAGAAAAA GATAATAACT CAGTTCATAT TTGCACCTAC 780  
 TTCAGTGGAC ACTGAAATTT GAAGGTGGAG GATTGTTT TTCTTTCTTA AGATCTGCG 840  
 ATCTTTGAA TCTACCTCT AATGATTAG AGACAGAGCT TGAGCTTAGC AGGGCAGAC 900  
 TTGTCCACC TGTTGTTCT TCTGACGAG ACTTTGAGCG TGTCAGAGCG CTTTGTGCT 960  
 GTGTGCTCCG GCAAGTTTCC TTCTCTGGAG CTTCGCCAG GTGGGAGCT AGCTCGGAG 1020  
 ACTACCGAT CATCAAGCC TGTGAACTC TTCTGACAA GAGAAGGGA GCGCGGGTAA 1080  
 GGGAGTAGG TGGAGATTG AGCCAAGTC AAGGATGGA GTGCAATTG AGCTGGGAA 1140  
 GGTATGCT CTGGCGCGT CCAAGACTA CCGAGAGCT TTCCGAATC TTTCGACGA 1200  
 GTGTGGGAAA GTGATCCAGA ACCCGGGCC CAGGACCCA GAGGCCGGA GCGCAGCAAC 1260  
 TTCGGGCGCC AGTTTGTCTG TGCTGACGA CAGCAGCAGC CAGCAGCAGC AGCAGCAGCA 1320  
 TGACGACGAG CAGCAGCAGC AGCAGCAGA GACTAGCCCC AGGCAGCAGC AGCAGCAGCA 1380  
 GGTGAGGAT GTTTCTCCCC AAGGCCATCG TAGAGGCCCC ACAGGCTAAC TGGTCTGGA 1440  
 TAGGAACAC CAAGCTTAC AGCTGCTGAG TGGCGGTGAG TGGCACCCG AGAGAGGTG 1500  
 CTTGCCGAAA CTGGAAGCC CGTGCGGCC CAGCAGAGGG CTCCCGGCGA AGCTGCGAGC 1560  
 ACTCTGGGAC GAGGATGACT CAGCTGCCCC ATCCAGTTGT TCCTGTCTGG GCCCCACTT 1620  
 CCGCGGCTTA AGCACTGCTC CCGCTGACCT TAAAGACAT CTCAGCGAGG CCAGCACTAT 1680  
 GCAACTCTT CAGCAACAGC AGCAGAGAGC AGTATCCGA GCGACAGCA GCGGAGAGCG 1740  
 GAGGAGAGCC TCGGGGGCTC CCACCTGCT CAAAGCAAT TACTATGGG GCACTCTGAC 1800  
 CATTCTGAG AAGCAAGAG AGTTGTGTGA GCGAGTGTG GTGTCTATGG GCGTGGGTGT 1860  
 GGAGGCGTGT GAGCATCTGA GTCCAGGGGA ACAGCTTCGG GGGGATGCA TGYACGCCCC 1920  
 ACTTTGGGA GTTCCACCCG CTGTGCTGTC CACTCTGTG GCCCAATTGG CGGAATGCAA 1980  
 AGTTTCTCTG CTAGACGCA GCGCAGGCAA GAGCACTGAA GATACCTGTG AGTATTCCCC 2040  
 TTTCAGGGA GTTACACCA AAGGCTAGA AGGCGAGAG CTCAGCTCTC CTGGCAGCGC 2100  
 TGACGAGGAG AGCTCCGGGA CACTTGAAT GCGCTTACAC CTGTCTCTCT ACAATGCCG 2160

ACGACTGGAC GAGGACGCTG CGTACCAGAG TGGGACTAC TACAACCTTC CACTGGCTCT 2220  
 GCGCGGACCG CGCGCCCTCC CGCGCGCTCC CCATCCGCCAC GCTGGCATCA AGCTGGAGAA 2230  
 CCGCTGGAC TACGGCAGCG CCGTGGGCGG TGGCGGCGG CAGTGGCGCT ATGGGGAGCT 2240  
 GCGGAGGCTC CATGGCGCGG GTGCAGGGGG ACCCGGTTCT GGGTCACCTC CAGCCGCGCC 2450  
 TTCTCATCC TGGCACAAC TCTTCACAGC CGAAGAAGGC CAGTGTGTGT GACCTGTGGT 2460  
 TGGTGTGGGG GTGTGTGGGG GCGCGGGGCG CGCGCGGCGG CGCGCGGCGG 2520  
 GCGCGCGCGG GAGCGGGGAG CTGTAGCGCC CTAGGGCTAC ACTCGGCGCC CTCAGGGGCT 2580  
 GCGGGGCGAC GAAAGCGACT TCACGCGACC TGATGTGTGG TACCCTGGCG GCATGGTGA 2640  
 CAGAGTGGCC TATCCGACT CCACITGTGT CAAAGCGGAA ATGGGCGCCT GGATGGATAG 2700  
 CTACTCGGGA CCTTACGGGG ACATGTGGTT GGAGACTGCC AGGGACCAAT TTTGGCCAT 2760  
 TGACTATTAG TTTCACCCCG AGAAGACCTG CCTGATCTGT GGAGATGAAG TCTGTGGTGT 2820  
 TGACTATGA GCTCTTGA TTGTAGAGTG CAAAGTCTTC TTCAAAGAG CCGGTGAAGG 2880  
 GAAACAGAAG TACTCTGGCG CCAGCAGAAA TGATTBCACT ATTGATAAT TCCTGAAGAA 2940  
 AAATGTGTCA TCTTGTGCT TTGGGAAATG TTATGAAGCA GGGATGACTC TGGGAGCCCG 3000  
 GAGGTGTGAG AAACITGGTA ATCTGAAACT ACAGGAGGAA GAGAGGGCTT CAGGACCAAC 3060  
 CAGCGCCACT GAGGAGACAA CCGAGAAGCT GACAGTGTCA CACATTGAAG GCTATGAATG 3120  
 TCAGCCCATC TTGTGTAATG TTCTGGAGGC CATTGAGCCA GTGTAGTGT GTGCTGACCA 3180  
 CGGACACAA CAGCGCCACT CTCTTGACG CTGTGCTCT AGCTCAATG AACCTGGAAG 3240  
 GAGACAGCTT GTACACGTGG TCAAGTGGGC CAAAGGCTTG CCGTGGCTTC GCAACTFACA 3300  
 CGTGGAGGAC CAGATGGCTT TCATTCAATG TCTCTGGATG GGGCTATGG TGTTTGCCAT 3360  
 GGGCTGGGGA TCTTCAACCA ATGTCAACTC CAGGTATGCT TACTTGGCCC CTGATCTGGT 3420  
 TTTCATGAG TACGCGATGC ACAAGTCCCG GATTTACAGC CAGTGTGCTC GAATGAGGCA 3480  
 CCTCTTCAA GAGTGTGAT GTCTCAAACT CACCGCCGAG GAATCTCTGT GCATGAAGC 3540  
 ACTGCTACTC TTACGATTA TTCCAGTGA TGGGCTGAAA AATCAAAAT TCTTTGATGA 3600  
 ACTTCGAATG AACATCACTA AGGAACCTGA TGTATCATC GCATGAAAAA GAAAAATCT 3660  
 CACATCTGTC TCAAGACGCT TCTACCAAGT CACCAAGCTC CTGAGTCCG TGCAAGCTAT 3720  
 TGGGAGAGAG CTGATCATG TACCTTTTGA CTTGCTAATC AAGTCAACGA TGTGTAGGCT 3780  
 GAGCTTGGG GATCTTGA TGG CAAAGATCT CTCTGAGTCT CAGTGTGGA TCCCTTGTG 3840  
 GAAAGTGAAG CCGATCTAT TTGACACACCA GTTAAAGCAT GGAAGACCTA TTTCGCCAC 3900  
 CAGGCTCATG CCGCCCTTCA GATGTCTCT CCGTGTATA ACTTGCAC ACTCTCTGC 3960  
 AGTGCGCTGG GGAATTTCT CTATTGATGT ACAGTCTGTC ATGAACATGT TCGTGAATTC 4020  
 TATTTGTGG GCTTTTTT TCTCTTCTC TCTTCTTT TCTCTTCTC TCCCTATCT 4080  
 AGCTCTTCCA TGGCAGCTC AGATGTGCT TCCATGTGTG GCCTCTGCT GTGTTTGA 4140  
 TGTGTTGTA TGGCTTTAAA TCTGTGATGA TCCATATGTT GCGCAGTGC AAGTGTGCT 4200  
 TGTTTACAGC ACTACTCTGT GCGACGACCA CAAAGCTTGA CTATCTTAT GCGACGGAAA 4260  
 GTTTGAGAG CTAAGATTAT CTGGGGAAAT CAAACAAAAA AACAAGCAAA CAAAAAATA 4320  
 A

Seq ID NO: 94 Protein sequence  
 Protein Accession #: NP\_000035.1

1 11 21 31 41 51  
 | | | | |  
 MEVQLGLRV YPRPSPKTYR GAFQNLFGSV REVQNPQPR HPEAASAPF GASLLIQQ 60  
 QQQQQQQQQ QQQQQQQQT SFRQQQQQG EDGSPQAHRR GPTGYLVLD EQPSPQPSA 120  
 LECHPERGCV PEPGAVAAS KGLPQLQPAF PEDDDSAAPS TLLSLGPTTF GLSSCSADLK 180  
 DILSEASTMQ LQQCQQAQV SBSSSSGRAR BASGAPTSK DNYLGTSTI SDNAKELCA 240  
 VSVSMGLGV EALEHSPGBQ LRQDCMYAPL LGVPPVPRPT PCAPLAECK SLLDSSAGKS 300  
 TEDIATYSPF KGYGYTGLEG ESLGCGSAA AGSSGTLELP STLSYKSGA LDFAAAYQSR 360  
 DYTPEFLA GPTPEFLA PHARILBNP LDYGSWAIAA AAGCYGDLA SLHGAAGAP 420  
 GSGSPSAAS SSWHTLPTAE EQLYGPQGG GGGGGGGGGG GGGGGGGGGG GGEAGAVAPY 480  
 YTPRPQGLA QGESDFTAPD VVYPGGMVSR VPPYSPCTVC SEMGPWMSY SPQYDMRL 540  
 TARDHVLPID YYPFQKTEC KGDDEASGCH YGALTGCSCK VFKRAABGK QKYLCASTRND 600  
 CTIDKPRKRN CPSCRLEKCY EAGMTLGARK LKKLGNLKLQ EGEASSTTS TPTETQKLT 660  
 VSHRGVQC PRLNVLBI EPVYVCAHED NNQPSRAAL LSLNBLER QLVRVYKVAK 720  
 ALGPFRNLIV DDQMAVQVS WMGLMVFAMG WRSFNVNRS MLVYAPDLVT NEYRNHKSRRM 780  
 VQVCRMRHL SQEFGWLQIT PQEFLCMKAL LFLSHVPDO LKNQKFFDL RMNYKELDR 840  
 IACKKKNPT SCRRRFYQLT KLLDSVQPIA RELHQFTFDL LKSDMSVSD PFMMAERS 900  
 VQVPRKLSK VKPIFYHTQ

Seq ID NO: 95 DNA sequence  
 Nucleic Acid Accession #: NM\_002497  
 Coding sequence: 135-1472 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 | | | | |  
 GGCACAGTA GGGGTGGCGG GTACATGTGT CTCGGGGGCT TCTCAATCA GGTCTCTGGA 60  
 GTTCTGGTCT CCGTGGAGCT CGACATGTGC GCGCAACCTG CGTGAGGCA CCGCACTCTG 120  
 GCGACTGGCC GGGCATGCTT TCCGGGGCTG AGGACTATGA AGTGTGTAC ACCATTGGCA 180  
 CARGCTCTCA CCGCGCTGCG CAGAAGATCC GGAAGGAAGG TGATGGCAAG ATATTAGTGT 240  
 GGAAGAAGCT TGACTATGCG TCACTACAGC AAGCTGAGAA AAGACTGACT GTTCTTGACA 300  
 TGAATTTGCT TGTGATGAG AAACATCCAA ACATGCTTGC TTACTATGAT CGGATTATGG 360  
 ACAGGACCAA TACACACTGT TACATGTGAA TGGAAATATG TGAAGGAGGG GATCTGGCTA 420  
 GTGTAAATAC AAGGGAAGC AAGGAAGGC AATACATTGA TGAAGATTT GTCTCTTGAG 480  
 TGATGACTCA GTTGACTCTG GCGCTGAAGG AATGCCACAG ACGAAGTAT GTGTGCTATA 540  
 CCGTATGCA TCGGGATCTT AAACCGGCA ATGTTTCTCT GGATGCGAAG CAAAGCTGCA 600  
 AGCTTGGA GA CTGGGCTA GCTAGATAT TAAACATGA CAGGAGTTT GCAAAACAT 660



TTGTTGGCAG ACCTTATTAC ATGTCTCTG AACAATGAA TCGCATGTC TACAATGAGA 720  
 AATCAGATAT CTGCTCATTG GGCTGCTTCG TGTATGAGTT ATGTGCAITA ATGCCCTCAC 789  
 AGACGCTTT TAGCCAGAAA GAACTCGCTG GGAATAATCAG AGAAGCGAAA TTCAGCGCAA 840  
 TTCCATACCG TCTCTCTGAT GAATTGAATG AAATTATTAC GAGGATGTTA AACTTAAAGG 900  
 ATTACCATCG ACTCTCTGTT GAAGAAATG TTGAGAAACCC TTATAAGCA GATTTGGTGG 960  
 CAGACGAGCA AAGAAGAGT CTGAGAGAA GAGCGGACCA ATTAGGAGAG CCAGAAATAAT 1020  
 CGCAGGATTC CAGCCCTGTA TTGAGTGAAG TGAACATGAA GGAAATTCAG TTACAGGAGC 1080  
 GAGAGCGAGC TCTCAAGCA AGAGAAGAAA GATTGGAGCA GAAAGAACAG GAGCTTGTGG 1140  
 TTGCTGAGAG ACTAGAGAG GACAAACTGG CTAGAGCAGA AATCTGTGT AAGAACTACA 1200  
 GCTTGGTAAA GGAACGGAAG TTCTGCTGTC TGGCAAGTAA TCGAGAACTT CTTAATCTTC 1260  
 CATCTCCAGT AATTAAAGAG AAGATTCAIT TCAATGGGGA AAGTAAAGAG ACACATGAGA 1320  
 GGAGTGGAAA TTGAGAAAGT CAGCTACAT CTAAAGTCAA GTSCAAGGAC CTGAAGAGAAA 1380  
 GCGTTCACGC TGCCGACGCT CGGCTCAAG CCTGTGAGA TATTGAGAAA AATTACCAC 1440  
 TGAAGAAGCA ACAGATCTCT GGCATCGCT AGCCAGGTAG AGAGACACAG AGCTGTGTAC 1500  
 AGAGTGTAA ATTACCAACC TTAAAGACT GATATACAAA TGCTGTAGTG TGAATATCT 1560  
 GGCCCATAGA GCAATGCCIT TCTGTATA GT ACACATGATA TTGCGAATTT GGTTTATCTG 1620  
 TCTTCAGCA ACTATGTGAC AAAATGTGCA CATTTAATIT TCTTCTCTC TTITTAAGAAC 1680  
 ATATTATAAA AAGAATACTT TCTTGTTGG GCTTTAATC CTGTGTGTGA TTAAGTAGAT 1740  
 GACATCAGAGA GTGACATTC TAAATCTTGG GAGAAAAAAT AATATTAGGA AAAAAATATT 1800  
 TATGCAAGAAA GAGTAACACT CACTGAATAG TTTTAAATGA CTGAGTGTGA TGCTTACAA 1860  
 TTGTCACTGT AGATTAAAT TTAAAGTCTG AGATTTAAA TGTTTTGA GCT CTTAGAAAAC 1920  
 CCAGTTAGAT GCAATTTGCT CATTAATAG ATGACATCT GCTTATAAAT ATTCCTATGC 1980  
 TCTGTAGTTC AAATCTGTTA GCTTTGTGAA AATTCATCAG TGTGATGTTT GTATCTTCT 2040  
 TTITTTCTGT TTTAACAGAA TATGAGCTGT CTGTCAITTA CTAATGCTT TCCACATGAA 2100  
 TAAAGAAGAT CTTCAGTTA

Seq ID NO: 96 Protein sequence  
 Protein Accession #: NP\_002488

1 11 21 31 41 51  
 GHRFNPIM ANEVRNIM TOSIGSENEA RELATEDKIN ASEIHOMOSAP IENSMPRAE 60  
 DYELTYGT GSYTKGIR AAKATGTGCA KRLDYSNTE AKKQALVSEY NLLRLKHPN 120  
 IVRYVDRID RINTLLYVM EYCRGGDLAS VTKGTERQ YLDFEFLVR MTQLTLAKE 180  
 CRRSSDGGHT VLHRDLKPA VFLDGKQNVK LQDLRLARL NHDTSFAKTF VGPYPYMSAE 240  
 QMNMESYNEK SDIWSLGLCL YELCALMPFF TAFSQKELAG KIRRGKPKRY PYRYSDELNE 300  
 ITRMLNLKD YHRPSVEEL ENPLJADLVA DQRGNLEIR GRQLGEPKRS QDSSPVLSEL 360  
 KLRKELQQR ERLAKRHRER LQKQZQLVCY RERLADKLK RAEHLNLSY LLKRSKRLD 420  
 ASNFLINLP SVSKIKLVKPS DESKEKNIMR SENSISQLTS KSKCKDLKKR LHAQLRLQA 480  
 LSDIEKNYQL KSRQLGMR

Seq ID NO: 97 DNA sequence  
 Nucleic Acid Accession #: NM\_007050.2  
 Coding sequence: 185-4576 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CCTCCCGCTT CAGTTCGCGC CGCGCTCGG CTGGAACGC AGGAGCGCC GCTCCGAGAG 60  
 CCCAGCGGA CGCAGCGCC CGCAGAGCA GCGCGCGCC CGCGATGCG GCGCAGCCC 120  
 GCGCCCGCCG CAGTCCCGCG CCGCGCGCC GCGGGAAGGG GCTGAGCTGC CCGCCCGCCG 180  
 CCGGATGGCG AGCTCTCGCG CGCTCGCCCT CAGCTGCTCT CTGAGGCTGC AGCTCGCGCC 240  
 CAGTCCCGCG GCGCGGCTC AGAGCGGCC AGGTGCTGT TCTTTGATG AGCACTACAG 300  
 CAACGTGTGT TATAGTGTGG CTCTAGGAGC CAATGGGTTT ACCTGGGAGC AGATTAAAC 360  
 AACGAGAGAA CCAATGCTGG ACCAGGCTTG GCGCAGAGTA TCTTTCATGA TGGTGAACAG 420  
 CTCTGGAGAG CTGAGTGTGG AGAAGGCTTC AGAAGTCCGA CTTCTCTCT CCAACCTGGA AGGAGAATGA 480  
 CACCCACTGC ATGCAGTCTC ATTACTACTT CTCACGCTGT GACAGGTCCA GCCCAGGGCG 540  
 CTGGAACGTC TACGTGAAGG TGAATGGTGG CCCCACAGG AACCTGTGT GGAATGTGTC 600  
 CGGGGTGCTC ATGCAAGGCT GGGTGAAGGC AGAGCTCGCC ATCAGCACTT TCTGGGCCA 660  
 TTTCATCATG GTGATATGCT AATTCGCTTC ATGAAGGCT CATCCGCTCT ACATGCGCT 720  
 GAGAGAGTGT GCTGCTGCTT GCTCAATGCA GAGAAGCA CTTCAATTT TGGAGTCCA 780  
 AAACGTGGAG GTGAATGTGG GCGCAAAATG CACATTTCAG TGCATTGCT GTTGGGAATG 840  
 GTCTCAGCAT GACAAGCTTT GGCTCCAGCA ATGGAATGGC AGGGACAGG CCTGTATGTT 900  
 CACCCGTGTG GTCAACCCCA GCGCTCTCT AGGCCACACT AGTGTGGCAG ACACGTGCCA 960  
 GCGGAGCTCT AGCAAGTACC GCTTGTGTAT CGCTCTGAT GTTGGTCTG GTGTGTCAA 1020  
 CTACGAGGAG CTGATGTGTA AAGAGCTTCC CAGCCCAAT GCTGCCCGAG AGCTGTGCG 1080  
 TTGTGGGGCC ACATACCTTT GGATCAGGCC AAATGCCAAC TCCATCATG GGGATGGCC 1140  
 CATCATCTTG AAGGAAGTGG AATATCGCAC CACCACAGGC AGTGTGGCAG AGCAACCAT 1200  
 AGTGCATCT CCAACATATA AGCTGTGGCA TCTGGACCCC GATTGTGAT ATGAGATCCG 1260  
 AGTGTCTCT ACACGACAGG ATGAGAGGGG TACGGAGCC CAGGGGCTC CCTTACACAC 1320  
 CAGGACACAG TGTGACAGAT CGGTACATGG CTACAGAGC GTGAAATGT TAGACATCAG 1380  
 AGCTCGGAG CAGACCTGCT AGTGGAGGCC CTGCGCTCAT GCGGTGAGCC GCTGCATCAG 1440  
 CTACACACT AGCTGTGCACT ACCAGTATGT GTTCAACAC CAGCATGAC AGGCAGAGGA 1500  
 GTTCACTCAG ACCCTCTCCD ACTACACCTT GCGAGGGCTG GCGCCCTTCA TGACCATCCG 1560  
 GTCTGCACTC TTGCTGTCTA ACCCGAGGG CCGAATGGAG AGCGAGGAGC TGGTGTGTCA 1620  
 GACTGAGGAA GACGTTCCAG GAGCTGTCTC TCTAGAGCT ATCCAAAGGCG GCGCTCTTGA 1680  
 GGGAAGATG TATATCTGCT GATGCTCCG CATGAGAGC AATGGGGTCA TCAAGCTTGA 1740  
 CGAGATCAAC TACAAGGCTG TCGGCTGCTT GAGCCCAAGT GCTGACCTT CAGGCGCAG 1800

- 5 GGGGAAAGTG TTCAAGCTCC GGAATGAAAC CCACACCCTC TTGTGGGTG TGTAACCCAG 1860  
 GACCACCTAT TCTTTCACCA TCAAGGCCAG CACAGCAAAG GGCTTTGGGC CCCCTGTAC 1920  
 CACTCGGATT GCCACAAAA TTTCAGTCC ATCTCATGCT GAATACGACA CAGACACCCC 1980  
 ATTTGAATGAG ACAGACADGA CCATCAACAT GATCTGTGAAA CCGCGTCAGT CCCGGGGAGC 2040  
 TCTGTCTAGT GTTATGACGC TGGTGTGAAA GGAGGGAGCGA CTTGAGAAAT CACGAGGGG 2100  
 AGCTGACAT ATTGAGAGCT TTTCGGTGCC CCGTAGCTAT CGEATATGCT CCAGCTGTGA 2160  
 TTCTCTACAC TACTTTGCTG CTGAGTTGAA GCTTGCCACG CTGGCTGTCA CCGACGCCAT 2220  
 TACAGTGGGT GACATAAAGA CATACAAATGG CTACTGGGAA CTCTCTCTCT CTCCCTGAAA 2280  
 10 AAGCTACAGC ATCTACCTCT AGGCCTACAG CAAAGAACAT GGAGAGACCA AATCAACTGT 2340  
 TGTCTCTCTG GCTACAAAAG CACCAATAGG CAACGCCGAG GTACCCCGGG GBACTCCACT 2400  
 CTGCTCTCTG ACAGAGGCTG CTTTACCGCA GALLTGTAC ACTGTGGAGC CAGAGAAGCA 2460  
 GGTGGACAC ACCGTGAAGA TGGCTGGGCT GATCCCTGGC CTCTCAITGT TCATATCAT 2520  
 TCTCTGGGCG GTAGTGCTCA CCATCAAAG GAGAAGAAAT GCTTATCTCT ACTCTATTA 2580  
 15 CTGTGCCAAA AGGAAGCTCG CCAAGAAAGCA GAAGGAGACC CAGAGTGGAG CCCAGAGGGGA 2640  
 GATGGGGGCT GTGGGCTCTG CCGACAAAAC CACACACCAAG CTCAGGGCGCA GCCGCAATGA 2700  
 TGAAGGCTCT TCTTCTAGT CTCAGGAGAT CTCAAGATCT ACAGATATGCA GCGCCGGGGA 2760  
 GCTTCCGAG CCAACATGCA GATCCGACG TCATCTGTAC CGACATGTG ACCTGTGTGA 2820  
 GATGAGCTAC CCGCGGAC ACCTTCCAAT CGCCATCGG GTGGCTGACT TCTCTGACGA 2880  
 CATCACGAG ATGAAGAGAG GCCAGGGCTA CGGGTTCAAG GAGGAATACG AGGCTTACC 2940  
 20 AGAAGGGGAG ACAGCTTGT GGGACACAGC CAAGGAGAT GAAAACCGCA ATAGAATGT 3000  
 ATATTGGAAC ATCATATCT ACAGCATTCT CCGGGTAGG CTGCTGTGTC TGGATGGGA 3060  
 GCGACGCTT GAGTATGACA ATGCCAACTA CATTTACGGA TACCATGAGC CTGGGACTA 3120  
 CATTTGCACT CAAGGTCGGA TGCAAGGAGC TGTAAGGAC TTTTGGAGAA GTATCTGGCA 3180  
 25 GGAGAACTCC GGCAGCATCG TCATGTGTAC AAACCTGGTG GAAGTGGGCA GGGTGAATGT 3240  
 TGTGGATATC TGCCAGATGT ACACGAGAGT CTAAGGAGAC ATTAAGTGCA CCTGATTTA 3300  
 AACAAGAGCC CTGGCAGAAAT ACCTCAATAG CACTCTACAC GTCCAGAGA AAGTGCATCA 3360  
 TGAGATCTCG GAGAGAGCTT CAGACGCTGT CAGGCTGAG GGGTGTGAG GTGAGCTCCG 3420  
 CTATGGCACT GGCCTTCTGG GCTTGTTCGG CCAAGTCAAG TTCTTCAAC CCCGGAAGC 3480  
 TGGGCCATA ATGTGTCAC CTGAGTGTGG GGGTGGGCGG ACTGGTCTCT TCATTGCCAT 3540  
 30 TGACACCAAT CTTCAGATGG CCGAGAATGA AGGGGTGGTG GACATCTTCA ACTUCGTGCG 3600  
 TGAGCTCCGG GCCCAAAGGG TCAACCTGGT ACAGACAGAG GAGCAATATG TTGTTGTGCA 3660  
 CGATGCCATC CTGGAAGGCT GCTCTCTGTT CAGCATCTCC ATCTCTGTGT GTAGTCTCTG 3720  
 TTCTCTTAC TACATATGCA GAGAGTGAGA CCGCAGACA AACTCCAGCC AATATCAAGA 3780  
 TGAATTTGAG ACCCTCAACA TTGTGACACC CCGTGTGGGG CCGAGGAGCT CGACATCTGT 3840  
 35 GGTCTCGGCC CGGAACCAAT ATAAGAAATG AAGTATGGAC GTGCTGCTC TGGACCGCTG 3900  
 CTGCGCTCTC CTATCTCTAG TGGAACGAGA ATCCAGACAT TACATCAAGG CACACATGAT 3960  
 GGAATAGCCAC AAGAGAGGCT GCTCTCTGTT GGTACCCAG CACCCTCTAC CCGAAGCT 4020  
 GCGAGACTCT TGGAGAGCTA TGTGATATTA CAACCTGTGA CTCTGTGTA TGTGAATGA 4080  
 GATGGACACT GGCACATCT GTATGCACTA CTGGCGTAG AAGACCTCGG GGTGCTATGT 4140  
 40 GCGATCCAG GTGAGTTGTG TCTGCGCAGA CATCGACGAG GACATCATCC ACAGAATAT 4200  
 CCGCATCTGT AACATGGCCC GGGCACAGGA TGGTATCTCT ATAGTCCAGC ACCTTCAGTA 4260  
 CATTTGGCTG CTGCTGCTAC GGGACACAGCC CCGCTCCAGG CBTCTCTGTC TCAAAAGTGT 4320  
 CGAAGACTCT GAGAACTGGG AGGAGCACTA TGACGGGAGG GAGGAGCTGA CTGTGTGCA 4380  
 CTGCTTAAT GGGGGAGCC GTATGGAAC CTTGTGTGCC ATGTGCACTG TTGTGTGAAT 4440  
 45 GATCCAGAG CAAAGACATA TTGACGTGTT CCACATCGTG AAACACATCG GTAACAAAA 4500  
 ATCCAACAT GTGGAGACCC TGGAACAGTA TAAATTTGTA TACGAAGTGG CACTGGAATA 4560  
 TTTAAGCTCC TTTAAGCTCA ATGGGATGGG GAACCTCGGG AGTCCAGAGG CTGCTGTGAC 4620  
 CAAAGCCCTCT TTGTGTGAAA TGGAGCTAAC TGGGCTCAGG AGCTCTGAGT TGGCACTCTG 4680  
 50 CTGTACTGCA ACAGAGAGAG AGCTGTGCGT GTGTCTACG GGGGCTGTGT CACTCTCTGA 4740  
 GGGGTCTCTT GTTGCGGTGG GAGATGCTGC TCCAAAAGGC CCAGGCTCTT TTITCAACCT 4800  
 AACCAGCCAC AGCCAAGGG CCAAGCAGAA GTACACCCAC AAGCAAGGCC TTGGATTCT 4860  
 GGGTCCGAGA CCACTGTCTT TGTGTTGAG TTGTGTGATC TTTGGCAAG CCAACTGTGC 4920  
 AGGTGCTGGG GAGTGGGAGG CTCCGCTGCC CTCTCTTCTC TTAGGAGTGG AGGAGATGTG 4980  
 55 TTGTCTGCTC CTCTAGCTGA TGGAAAGAT TGAAGCTGTT GGGGTGCTAC GTCTCTGTC 5040  
 CCGGTGCTAC CTCTTCAGG GGGCTCTGGC ACCAGACAT ATGACATGCG ACCAGATGTGA 5100  
 CCTTACATGT TTCTTAGGC CACAAGAGAG GCGGCCCAT CTCACACCTA ACTCGATGCG 5160  
 GGCCTGCCCC ACAACATCT TGATCCCTT CCGCAGGCTG GGCCTTGACC GTCCAGACT 5220  
 CACTGGCGCG CCACTGTGTT GACACGACAT TTITGATAAA GGTGTGTTCT GCTTTTGT 5280  
 60 GTGTCTGCTG TGAGGGGTGT CCACTGATGG GAACTCTCT GCTGCTCTCT GCTCTGTGAA 5340  
 AAAGGAGCTA CAGTCTGAG GCGGCTGTGG ACCAGACAT ATGACATGCG ACCAGATGTGA 5400  
 TCTTGTGTGG TACGTTCTT TGGAGAGAGC AAACAAGATA AGATTGATAT ATTTCZAAA 5460  
 GGTATGTGGA AAAGAACCTT TCTTTGGAG GTGTGATAAT CTTAGTCTCT TATGTCAAAA 5520  
 AGAAGGGGCG GGGGGAGGTT GAGTATGATC CTCTAAGACA AATCTCTCGG GCGTTTATAT 5580  
 65 TTCTCTGGC AATGTCTTTA AAGGCTCCCA CCGTGGGACA GATGCCACT GACCAAGGAG 5640  
 AGATGGTGA GCTTAAAGAT GTGCTCTGTG GTGCTCTGTG CAAATAGAG ACCAGTTTG 5700  
 TGCTAATTT GGAATGTCAA ATTTGAACTC CTCTCTAAAG AAACCCAGCA GCCACTTGA 5760  
 AAAAGGCCAT TGTGGAGCCC ATTAATCTTT GATTTAAAT AGGCCAAGG AACTCAGGCT 5820  
 GGAGATGAG GTGCTTGTCC AAAGTGTGAG TGAGTCAATG AGAGGGAACC AACATTGCT 5880  
 70 AAGTCTTAC TGTATGCGAG GGATCATGCT TGGCATCTC CATAGGACAT TTCAACAGT 5940  
 CCTTAGAAC CCCAGGAGAG AGCTACTGAC TTGTTATCAT CTCTATTGA TCATCTCTG 6000  
 CAATGAGAA ACCACACAC CTTCCTTAGT AATGAATCT TGGGTCTCAA AGGGCGAGGT 6060  
 AATGCAATG AGACTCTCC GTGCTGTTTT CTTCATCTC TCTAAGCAA GCAATTTT 6120  
 TATGGAGGGA AAATAAGGCC AGAAACTCTT GAGCAGATAA CTCACAAAT GGAAATTTAG 6180  
 75 TACTTCTCT CTGATGCGAG TCTCTTGGG AAGGBCAGAA TTCAAGATAT ATTTTGAAT 6240  
 CACATTTCCA GCTCCCGAGG AAAGCCAGCT TCACTTAAAT TCTTAGTCA TAAAGAAAT 6300  
 TCCCTTCTG TTGAGGCTAT GATGAGTACA GCGAGGAA CTCTGACCT TGATCTCTGA 6360  
 CCACTGCTTA GGGCATAAT CTGACAGCT CAGTGGGCT GGGACCTAGG AAGTCCATC 6420

TTAGAGGCTG GTCTAGCCDC AGACAGGGCA TGAGGGGGCAG AGAATTCAAG AAGGTACAGC 6480  
 TTGGCCCTC AAGAGCCCACT TGTATGCTGG GGAAATGGAA CCAATGGTCA GTATGTGGGA 6540  
 5 GTGGAATGAG GTTCCATGAG CTAAGAGACA AGAAAGCTCT TCGGCTCGC GGCTCTCGG 6600  
 AGAAGGGGAC GTCCATTCAT CTGGGGTCTT AACAAAGCAT AAAAGGAAA AAAAGAAACT 6660  
 CAGCGAAAGG GATCCATATG TGCATATGCA AAGAAATGTG AAAAGGCAT GTGAGAAAGCA 6720  
 GTCTGGTACA GCGCAGCCA GTGAGGCGAC AGCACAACAC GGGAGCGACG AAGAGATGAG 6780  
 CCAGGCTCCA GGAGACAGAT GCCATGCGG AGTACAGACT TTGTCTTATT GGCACAAGAG 6840  
 AGTCCATGGA GCTTTAGAGA GATGCACTCA GCTTGTGTGT GGCCAAAGCT CCTCTGGGC 6900  
 10 CAATGGGGCT CCTCTTTCT CTTCATCAG ACACATGTGA AACATTCCTT TAAGCGTGCA 6960  
 CTTTTTAAAT TCAATCATAT TTGTGTGTGT GCTCATTTGT TTGTGTGTGG AACATAATAT 7020  
 CAGGTTCAGC AAGCAATGAG TCGAAGGCCA TGAGAGATGGA ACATGAAAGG AGCCACAA 7080  
 AGTTGCTTCC ATTGGCATTG GCTCTGGAGC GTGCCAAGT TCCAGGAAG TCAAGGGA 7140  
 15 TCAAGGAAGG GCTGTCACTT TAGAGGTTC AAGGAAGTG CCTCAAAGCA AAGGCAAGCA 7200  
 AAGGAACCCC AGCATGAAT TGTCTTTTCT GCTCATGTGG CCTCTCCCA GTGTATTTC 7260  
 AAGCAGACCC GGGGACCCAC CCCCATCGGG CTCGTGTGCT TCGTCTGGCT CAGCCCAAT 7320  
 GCGCAGCTG GCTCTCCCA GCTCTCAAG AGCCTCTAGC ATGCAAACT TCGTCTGT 7380  
 20 ATGCTATTTT CTTAGATCTT GGTACATCCA GACAGGATGA GGTGUGAGG AGAGCTATT 7500  
 AACACAACT CTAAGATT 7560  
 AGACAGTGGC TTTTATCAT TTAATGGTA GGAATTTAAG GTGTGACTT AGGAGAAAC 7620  
 AAACCTGCAA AAAAATAA TCTCAGGCCA TGTGGGGTA ACCCAGCAAG GCGCCATGAT 7680  
 GTATATCCC AGTCAATCC TATTCTTTC CAZAAAGCA CCAATTCCTA AAGCAGAGA 7740  
 25 GTGATAGGT CTTAGGCGAG TGCACACAG AAGAATATGA GCTTATGGA TGGGGATGAC 7800  
 TTCCCTAAGA TCCCATGGGA CAAGGATGTT GCAAGGCTTG GATGAGATGG GGCACCAAT 7860  
 CCCAGGAAT TGAACATTTT CTTTAOCCA GGAATCTCC GGAGCCAACA CCAACCCCTC 7920  
 CAGGGGGTCT CCCCACCCA CCCCATTAC AGGGTGGCT CAGCCTGTCA TGAGCAGAGG 7980  
 AAAAATATTAT TAATGCTCT TGAGTCTTCA CACAGAGAG CTTCACCTCA TGAATGTGG 8040  
 CTCCTTTGG GGAATGCTA AGAAGTGA GAGAAAGCCA GGAATTTCT TCGCTGTGT 8100  
 30 TTATGGCTTA AATAGCTTCA GGTATGATCT TAGCTGCACT CCAAGATTGC ATCTCTTG 8160  
 GGGTGAAGAA TCTGGGCCAA CCAAGGGGCTC TTGGGGCTCT AGAAGGCCAC AGTAGGCCCT 8220  
 TCTTTTGGG AATGGAAGG GACAATTTTC TTTTAGTCT GCGCCCTCT GTGGGTGTGG 8280  
 CTTGCAAGG AACCAACAGA CCTTATGTC GGGACATCTA CATGTGAGCT CATTAATTT 8340  
 35 TCCAGCTT CTAATGAGGA GTCTGGAAT TGTGATGAG TGTGATGAG AACTAAGC 8400  
 TCTAGGGGA GAAATCACTT GGCACAGTT CCAAGCTAG TGAGTGAATG AACCAAGAT 8460  
 TAAACCGGTT TTTTCTACT ACAGAGACAA TATTTTCCA CCAATGTATC TCACATTTT 8520  
 CCCAGGAGGT TACCATGTA AGTATGAGT AGAGTGAAC AGATACGCTA GTGTATAAG 8580  
 CTCAAAGCAA ACAACAGTAA GCTTAAATAT CTTCAATGAT CTTATGTTT ACCTTGCAA 8640  
 40 TCTAGTCAA ATTGCTACT CACTTATGGA TTGACCTTGT TTGTTTAA TATGACTA 8700  
 TGAATATTAT AAGAAATGA GTGCTGTGT CTTGATGTC TTCTTTCTG TTACCCCGC 8760  
 1YTAGCGAAC CTAAGTCACT CTGTGCTTGA GCTTGACCT AGAATTGACT CTAGAGGAC 8820  
 TGACCTGTCT GCTCCGACA GCGAGTATA GGTCAAGAT CAAGACCAAC TGACCTTCT 8880  
 CTAGGAGCT CTTTGTGT GTGGGTGCT TGACCTACT GTTATGAGG GGACCTGAAC 8940  
 45 TAGGCACTT TCAAGTGGG TCTGGGAAG AACCCATAA CTCACACTAG AATGATGAG 9000  
 ATTGCTCT CTGGGTGGA GAGGATGAG CCCAAGAAC CTAAGAGGA AAGAGAGAC 9060  
 TGGACACAGC TGTACTCAG AGATTCTTGA ATGCTAGCT GGAAGTGTG GCTGTGTCT 9120  
 CAGTGGAGT CACATGTTTG CTAATGTGG CAACTCTGAG GACACACTT ATGAGCACT 9180  
 GGGGTCTGGA AGGCTCCTCA CTTAACCTA GGCACATA ATTACTGGT GCTCAGACA 9240  
 50 CTTAGCACT TGAGGGGGG ACTATTAGA AATCGAAT ATGATGAGC AATTAATTC 9300  
 TGGGTATGA CTGGGTGTT GTGGGACAGA GTCACCTCT TAGTGTGAC GAAAGAACT 9360  
 ATACATGAA AACATACATA AGTCTGATG AAGAAAGAC CTTCACTAC ATTAGAAGG GTAGATTGG 9420  
 AGCATTTAG GCGAGAAAC AGGAAGCGAA GCGCAGGAAA CTGGAACCCA GTGAATACT 9480  
 AGAACCGAGG ATGCAATGA CTTATTAGC AAAATGTACA CTTCTGTGAC ATAGCTGGAG 9540  
 55 AAGAGATGG TACAAGCTG TCCAGGCCAC TTGGAACAG GGCAAATCTC ATGTGTCTGG 9600  
 GCAAAGATGG ATGCAATTTT CTTAGCCA TCAATGTCT CATTCTTCA CTCCTGAC 9660  
 CCAACATGAT TCTGTGTG ATGTCTGAC AGGATACCA TTTCCTGCT CGTGTGGGA 9720  
 ACCCTTAGA GTACATAGAG GTATCAGTCC AGTAAAGCTT CTTACACAA CAGAAGTGG 9780  
 GCGCAGGAG TAGCAGCCAG GGCCTTATCT TGTACTCT CTAGGAGTGA CTCGCCAAC 9840  
 60 CAGATCCGA GACATTTAG GAAATGATA TTCTTGTGA CTTCACTGCC GTTGGACAA 9900  
 ATGAGAAAG ACACCTCTT TTAGCTTGA TATCTAGGCT CTTGGGCTGG GTAAAGACT 9960  
 CATGACACC AGGTCAACC AGGATATCA TCTGGGAAGA GATGGGTGAG CCGAAGAGA 10020  
 AGGTCAACA GAGCGCTTT CTAAGTCACT GTACACAGGT TGTCTGTGGC CAGATCTGC 10080  
 AGGGAATAAA ATGTGCAAC CAACGTGGAG GTAGTTTGA AGAATAAATA CATAAAGCTG 10140  
 GATAGTCATG TCACTCAAA TCCCTGACTT GGCCTCCCA TTACTTGACA GTCTGAGCT 10200  
 65 CTTCTAGCC TGTGACAGC TTCAAATCAC AGCCAGTAA AACAAAGAAA TAGGAAAGAT 10260  
 AAATCCACT AGAAGAGACA AGCTGAGATT CAGATTGTT TACTCTCC ATGTCAAAT 10320  
 TCTGTGTGG AGGTTTCCA TGTATACATG TCTAGAGTGA ATAGAATGA AGAGCTTGC 10380  
 1TTCTTCTG AGGGATCTGC CTTGAGTCT ATAGACTGA CAGCAGGAG AGGCTGTAT 10440  
 GGTGAGTGT GGGGGGAGCT GTTCTAGCT CAGTTTCTC TGACACATT TTAGGATCA 10500  
 70 TGATCTGTAT CTTCCGAAG ACACAGAGCA TATCTAAGCT ATATTGTG ACATGAGCAG 10560  
 ACTCTCTAG TTTTATGTA ACCAGGAGT GCGCTTTGCA TGCACTGACT ATAGAGATG 10620  
 TCTTATAG ATCAGAGGAG TCTTTTCTG CACTCCAGA AAGATAGTGA 10680  
 AAAGTCTAC AAAGGGCAT CACCAACTGA AATCCACTA TGAATGTG CTCTTAA 10740  
 GGAGGCATCA ACATCTCAA TGGTAGCTC CAAACCTAG ATCCCACTA TCTAAGACT 10800  
 GAGGGTGTGT CCACTGTGGC AGATACAAGG GAAAGTGAAG GCTCAGGAT GAAAGAAAT 10860  
 75 CTTATGGAA GAGTITTAG GGTGTATCA TTAATGGGCT TCTCTTATA TCTGAGACT 10920  
 CTTGATGTG GTGAGTGTG GACTGTAGAT CTTATGGA ATTTGGAG AATGAGTTC 10980  
 TGTGGGCTT GAGTGTGG ACAGAAATAT ATCATATAA AAGCTTGGAG CTCAGGTTAA 11040

TTAATGTAGT TCATGGTCTC TTATGTAGCA GGACTCTTGG ATGTGGAGGA GAAAGGGTCA 11100  
 TAGGAAGTAA ACCACCAAAA TTACAAAATT GAGTCTCTGT ACAAATCTCT CAGTGCCCTT 11160  
 GGCGCTATGA ATACAAATCA GTGGGCGCTC TCTATGATGG TCCAAACAAC TCTAGTGTG 11200  
 CACCGCTGCC CTGTATCTCC CATGGGAAGAT GAATAATGTC AGGTGTCTCT TGGGTCAAAG 11280  
 GCGCCAGGCG AATGTGGAGC GTTAGAGGCG AGAGTGGTGT CATTCACAGT AAGGTAGGCG 11340  
 TCTTAGAGGG TCAGGACAGAA TATGGTGTCC ATATCTTCCA TAGCTCTGCA GATCTTGGAA 11400  
 TGAAGTCAAG CACAGTTTGC TAGACCAAGG TCACCTCTCT GAGTATAACT AGGACCCATT 11460  
 AGTGAAACCT AATAGCTGTA AGGAAGAACC TGCTGTCTGC CAGAGAGGAT AAGCTGCCCA 11520  
 TCTCAGCAGC TGCTTAAAG AGGCGCAGTG TCTCTTAAAA GGGAAAGAGAA GCATTTGGTGA 11580  
 AATGGATCTT AGGTCACTTC CATTCAGAT GGGTAGAGAT GTTGGAGAGCT GGGATGATGT 11640  
 TTGAAGCTAT TCATACCTGT AAGACAGCAA TCCAGATGAA TTGTGTGTGG TCTGTACAGG 11700  
 CTGAAGGCCG CTGCTCTCC ACCCAAGTGC CCCCACGTAG CAGGCCAACA TGCTGTGTGT 11760  
 GCCACATATA CTGGGCTGAT CCAGGCTGGT TATCCACAAA CAGCAAAACA TAGGGAACAG 11820  
 CTGCTTGGCC ATAGACCCAA TAOCATGTA GATCTCTCAT GAGGAGCAGC ATAATCTAGA 11880  
 CCCCAGTACC AACAGGGCCA TGAGTGACAB CCAAGAACCA TGAAGGTGCA AGTAGGACAC 11940  
 AGAGCAGGCT TTTCTTACC ATACACATTA TCTCCAGAGG TTATTCTAC CCCACTCCT 12000  
 ATTCAAGGCC GTTGTGAGCA CACTGCCAAA GCAAAAGCAC AGTAATCTCA TTACACATG 12060  
 ATTATAATCA TTTCACGTGC ACACATTCCA TACCACGGTG GATCTCTGAGC TAGGCCATGT 12120  
 AAATCCGGGT TAAOCCATAT TGGTAATCAT ACTCAAAAGC ACTTTTCACC CTACATCTTA 12180  
 CTAGCCCAATC AAGACACAAG AGTTGTGGGC TCTACCATTO CTTGTGGCTC TGAGACAGCT 12240  
 CACAGATGAT CCAATGCTTC CCGGTCACT CCAGGAGGCG TGACATCTTC ACATCCAGTG 12300  
 GGCAATATAT ATTGCATGAG ACCAANGTCT CCACAGCTCT TGACGCTCC TCACTGAATC 12360  
 CCAATGGCCT GCACCTGTAC AGTTTGGGTG TTGTAGATAG AAGCAGTGA TGAAGAGAGA 12420  
 AAACAAAAATA AATCAACTTT TTAATAAAGC CAGCAGCTGT CTGTCAATGT TTTTCTTCT 12480  
 TTTCATATTC TAGCTCAGAA AAGCAGAAAG TAATAATGT CAGGTCAATG AATATCAGAT 12540  
 ATATTTTGT ACTGTACATT ACAGTGAAGT GTATCTTCT TACACCTGCA AGTCCATCT 12600  
 ATTATCTCT GTAAATGCT CCTGACAATG TTGTAAATG GGTCTGTGTA AAAATCATAT 12660  
 ACAATAAGC TGTGACCTG

Seq ID NO: 98 Protein sequence

Protein Accession #: NP\_008981.1

1 11 21 31 41 51  
 GIREFNPFRO TINTYROSIN NPHOSPHAT ASERBCEPTO RYTPETHOMO SAPIENSMAS 60  
 LAALALSLLL RLQLPLPQA RAQAPGGCS FDEHYSNQY SVALGTNGFT WQINQTHEP 120  
 MLQAVPTES FMHNMNRA SQKAEILLI TLKRNDECI RYFVSSSD RSGPALNAY 180  
 PVVNGQPNQV PVWNSGVT EEWVKEALAI STFWPEHYQ IFESVSKHG PGVAYADVR 240  
 VLAHFCRCAQ HFLRLQNVVE NVGQNAITQC IAGRGWSQHD KLWLQQWNGR DTLAMVTRVV 300  
 NHRFSAITVS VADTARQSVS KYRCVRSQG GSCVSNNV AELIVKBPPTIA PELLAAGAT 360  
 YLWIKPNANS IGDOPHLK EYEVKRTTGT WARTHVDSP NYKLWHLDFD VHYEIRVLL 420  
 RFGGGTGGP GAPTTRTKC ADPVHQPQW BIVDIRARQL TLQWEPFOYA VTRCISYNLT 480  
 VOYQYVPAQK QYBAEIVGT SSIITLRLGL RPFATRLLL LSNBQRMGES EELVQTEED 540  
 VFGAVPLESI QCGFEEKLY IQWFKFNETN GVITLYEINY KAVGLDFPSA DLSQQGKVF 600  
 KLURNETHLFL VGLYPTTYS PTKASTAKG PGPVPTTRIA TKISAPSMPE YDITDPLNET 660  
 DTTITVMLPK AQSRGAPVS YQLVYKEERL QKSRRAADI ECFSPVPSYR NASSLDSLHY 720  
 FAABELKANL PYQTFTVID NKTNYGWNF PLSPKYSYI YFQALSKANO ETKDNCVRLA 780  
 TKAPMSAQY TDTFTCLLT TASTQNSNT VEPEKQVNT VKAGAGVIAGL LMTILLGV 840  
 MLTKRRRNA YSYSYLLSK QKLAKQKEITQ SQAQNMGPV ASADKPTTKL SARNDDEFS 900  
 SSKQDVNGFT DGSRLGSQL PTLTQTHPYR TCDPVMSYF RDQFLAIRV ADLLQHTQM 960  
 KRQGGYGFKE EYEALPEGQT ASWDTAKED NKNKNYQNI ISYDIHSRVL LVLDGDPHSD 1020  
 YINANYIDGY HRPRHYIATQ GPMAQETVKDF WRMIWQENSA SIMVYTNLVE VGRVKCYRYW 1080  
 PDDTFYVYGI KYVLTIELPL AGYVIRFTPV QKQVYHEIRE LRLHFESWPF DHPVPCYAT 1140  
 LLGFVRQVF LNPISAGPIV VICSAGAGCT GCHADITML DMAENESVVD IFNCGVRLKA 1200  
 QRVLVQTER QYVEVHDAIL EACLGNATP PVCEFRSLYY NISRLDQTN SSQKDEQFT 1260  
 LNIPTVRVRP EDCSGLLPR NHDKNRSMVD LPLDRLCPLI ISVDGESSNY INAALMSHK 1320  
 QAAAPVITQH PLPNTVADF RLVEDYNCS VVLMNEMDTA QFCMQYWEK TSGCYQPIQV 1380  
 EFPVSADDED IHRFRICN MARQDQGYRI VQHLQYIOWP AYRDTYPSKR SLIKVVRLEL 1440  
 KRWQSQYDRE GRVYMEIGLN GGRSGTHCAI CSCVMEIQQQ NIDVFLIKV TLRNKNNSMV 1500  
 ETLQYKPYV EYALEYLSF

Seq ID NO: 99 DNA sequence

Nucleotide Accession #: NM\_002988.1

Coding sequence: 71-340 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CCGGACACAG AGGAGTTGTG AGTTTCCAA GCGGACGTCA CTCTGACCAC TCTCTGCT 60  
 GCGCAGCATC ATGAAGGCGG TTGCACTGCG CTCTCTGTCT CTCTGTGTCA CCAATGGCT 120  
 CTGCTCTGTG GCGAAGAGTG GTGACACAAA AGAGCTCTGC TGCTCTGTCT ATACTCTGT 180  
 CGAGATTCCA CAAAAGT TCA TAGTTGACTA TTCTGAACAC AGGCCCAAGT GCCCAAGCG 240  
 AGGTGTATCT CTCTCAACCA AGAGAGGCGG CGAGATCTGT GCTGACCCCA ATAAGAAGTG 300  
 GCGCCAGAAA AATACATCA GCG ACGTGAAGCT GAA TGCCGTA GGGGCTGGA AGTCGQAGG 360  
 GCGAGCTGAA CTGDTGGCG CCGAGAGGGA ACAGAGGCTT GAGCCAGGCG ATATGCTGCT 420  
 CCACCCGCA GCGCAGCTCT TGTAGAGCTT CCAATGTCTA TCCGACGA GATTAATCA 480  
 CTATTAATCT AGTTATGCA TCAATTCCA TTGTGAATAT GATTCTCAT GTTGGAGCT 540  
 ATTATGAAT TAGTATTTCT TCTGACATCT CATGCAATGT TCTTTATCAT CTTTGGCT 600

TTCCCTTCAA CTCTTCGAT ATTCAATGCA TGGATCAATC AGTGTGATTA GCTTCTCAG 660  
CAGACATTGT GCCATTGTGA TCAAAATGACA AATCTTTATT GAATGGTATT GCTCAGCAC 720  
ACCTTTTAAT ATATTGGCAG TACTTATTAT ATAAAAGGTA AACCGACATT CTCACTGTGA  
AAAAA AAAAAA AAAA

Seq ID NO: 100 Protein sequence  
Protein Accession #: NP\_002979.1

1 11 21 31 41 51  
| | | | |  
MKGLAALLV LVTMALCSQ AQVGTNKELC CLVYTSWQIP QKFIYDYSET SPQCPKPGVI 60  
LLTKRGQIC ADPNKKWVQK YISDLKINA

Seq ID NO: 101 DNA sequence  
Nucleotide Accession #: NM\_015507.2  
Coding sequence: 241-1502 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
| | | | |  
CCGCAGAGGA GCTCTGGCCA GGCTAGCCAG GGGGCCCCCA GCGCCCTCCC AGGCCGCGAG 60  
CACCCTCCGC GCGGTCTGCT GCTCTCCCTC CCAGACATGCA GGGACAGCAC CGGTGAACGT 120  
CCAGTGGAGC GGAGGACCGG AGCGCTGTAG GAGAGAGGAG GCGGCGGCTT ACCTGCTACG 180  
GGGTCCGGCC GGCGCCCTCT CGAAGGGGKC TCAGGAGGAG GAAGGAGGAC CCGTGCGAGA 240  
ATGCTCTGCT CCGTGGAGCTT TGGCTCCCGC CTGCTGCTCT CCGTGGTGCC AGGTGGTTTC 300  
GGGAAAGCGG CCAGTGCAGG GCATCACGGG TTGTATGATC CGGCACGTCA GCTTGGGGTC 360  
TTTCACTATG GAACATAACT GCGCTGCTGC TACGGCTGGA GAAGAAACAG CAAGGAGATC 420  
TTGTAAAGCA CATGAAACCA TGGTGAAG TTGTGTGAGT GCGTGGGACC AAACAALATG 480  
AGATGCTTTC CAGATAACAC CCGGAAAACC TGCAGTCAAG ATGTGAATGA GTGTGGAATG 540  
AAACCCCGGC CATGCAACA CAGATGTGTT AATACACACG GAAGCTACAA GTGCTTTTTCG 600  
CTCAGTGGCC ACATGCTCAT GCCAGATGCT ACGTGTGTGA ACTTAAGGAC ATGTGGCATG 660  
TATAAATGTC AGTACAGCTG TGAAGACACA GAAGAAGGHC CACAGTGCCT GTGTGCATCC 720  
TCAGGACTCC GCTGCGCCC AAATGGAGA GAAGCTCTAG ATATTGATGA ATGTGGCTCT 780  
GGTAAAGTCA TCTGTCCCTA CAATCGAAGA TTGTGGAACA CATTGGAAG CTACTACTEC 840  
AAATGTGACA TTGTTTTCGA ACTGCAATAT ATCACTGGAG GATATGACTG TATAGATATA 900  
AATGAATGTA CTATGGATAG CCAATACGTG AGCCACCATG CCAATTGCTT CAATACCCAA 960  
GGGTCTTCA AGTTGTAAGT CAAGCAGGGA TATAAAGGCA ATGGAGTCTG GTGTTCTGCT 1020  
ATCTCTTAAA ATTCTGTGAA GGAAGTCTTC AGAGCACTGT GTACCATCAA AGACAGAAATC 1080  
AAGAATTTC TCATACAA AAACAGCATG AAAAAGGAAG CAAAAATTA AATATTACC 1140  
CCAGAACCACA CCAGGACTCC TACCCTAAG GTGAACCTGC AGCCCTTCAA CTATGAAGAG 1200  
ATAGTTCGA GAGGCGGAAA CTCTCATGGA GGTAAAAAAA GGAATGAAGA GAAATGAAGA 1260  
GAGGGGCTTG AGGATAGAAA AAGAGAAGAG AAAGCCCTGA AGAATGACAT AGAGGAGCGA 1320  
AGCCTCGGAG GAGATGTGTT TTTCCTTAAG GTGAATGAAG CAGGTGAAAT CCGCTGATT 1380  
CTGTGTCAAA GGAAACGCTT AATGTCAAA CTGGAACATA AAGATTGAAA TATCTGGTT 1440  
GACTCGAGCT TCAATCATGG GATCTGTGAC TGGAAACAAG ATAGAGAAGA TGATTTTGAC 1500  
TGGAACTCTG CTGATGAGAA TAATGCTATT GCGTCTTATA TGGCAGTCC GCGCTGTGCA 1560  
GTCTCAAGAA AAGACATTGG CCAATTGAAA CTCTCTCTAC CTGAGCTGCA ACCCCAAGC 1620  
AAGCTTGTGT TGCTCTTTGA TACCGGCTG GCGGAGAGCA AAGTGGGAAA ACTTCGAGTG 1680  
TTTGTGAAAA ACAGTATGCA TCGCTTGCA TGGAGAGAAG CACAGAGTGA GATGAAAA 1740  
TGGAGACAG GAAAAATTA GTTGTATCAA GGAAGTGAAG CTACCAAAG CATCATTTTT 1800  
GAAGCAGAAC GTGGCAAGG CAAAACCGGC GAAATGCGAG TGGATGGCGT CTGCTGTGTT 1860  
TCAGGCTTAT GTCCAGATAG CTTTATATCT GTGATGACTG GAATGTTACT ATCTTATAT 1920  
TTGACTTTGT ATGTCAGTTC CTGCTTTTTT TTGATATGTC ATCATAGGAC CTCTGGCAT 1980  
TTAGAAATAC TACGTGAAAA ATTGTAATGT ACCAACAGAA ATATTATGTT AAGATGCTTT 2040  
TCTTGATATA GATGTCGCAA TATTTGTTT AATATGATA TCACGTATAT CTCTGATGCA 2100  
TTTCTGAATC TTTCACATT ATATTAAAA ATATGAAAAA TGCAGTTTAT CTCCCTCTC 2160  
CATGATATCT GATTTGTATA AGTAAATGGA TGAGCTTCTC TCTACACAT TCTAGAAAA 2220  
TAGAAAAAAA AGCAGACAGA AATGTTTAACT TGTGTGACT TTATGATCT TCTGGAAGAC 2280  
TATGACATCA AAGATAGACT TTTCCTAAG TGGTGTAGCT GGGTCTTCA TAGCAAAACT 2340  
TGATATTTA AATCTTGTG AATAATAATA TCCAATCAT CAAAAA AAAA

Seq ID NO: 102 Protein sequence  
Protein Accession #: NP\_056322.2

1 11 21 31 41 51  
| | | | |  
MPLPWSIALP LLLSVVAGF GNAASARHHG LLASARQPQV CHYGTXIACC YGWRNRNKG 60  
CATCEBQCK FQBCVGNKC KQFPYGTOK CSQDVNBQGM KRPQQRHCV NTHGSKYKFC 120  
LSGHMLMPDA TCNNSITCAM INCQYSCBTD BEGQQLCPQS SGLRIAPNRP DCLDIDRCAS 180  
QKVICPNNRQ CVNTHGYSYK KCHGPELQY ISGRYDCEID NECMDISHIT SHHANCPTNQ 240  
GSKFKCKGQY VYVNGSLRCA IPENSVKRVL KAPQTEKRI KILLAHKNSM KKKAKINVT 300  
PEFTKTPK VNLQPFNYVE IVSRGNSNGS GKKGNIEKMK BULDEKREK KALKNDIEHR 360  
SLRGDVFYFK VNEAGFGLI LVQRKALTSK LEHKDINISV DCSFNHGICD WKQDREDFED 420  
WNPADRNAL GFYMAVPALA GHKKDIDRLK LLLPLDQCS NFCLFDYRL AGDKVGKLRV 480  
FVKNSNNALA WEKTSIDEX WKTGKQLVQ GTDATKSIH EABRKQKGTQ ELAVDGLVLV 540  
SGLCPDLSLLS VDD

Seq ID NO: 103 DNA sequence

Nucleic Acid Accession #: NM\_001565.1  
 Coding sequence: 67-363 (underlined sequences correspond to start and stop codons)

5 1 11 21 31 41 51  
 GAGCAATTC TCAATTCCTT AGACATAATC TGAGCTGTACA GCAGAGGAAC CTCACGTCTC 69  
 AGCAACATGA ATCAAACTGC GATTCGTGAT TCGTGCCTTA TCTTTCTGAC TCTAAATGGC 120  
 ATTCAGAGAG TACCTCTCTC TAGAACCGTA GCGTGTACCT GCATCAGCAT TAGTAATCAA 180  
 CCGTGTAAAT CAAGGCTGCTT AGAAAACTCT GAATATATTC CTGCAAGCCA ATTTTGTCGA 240  
 CCGTGTGAGA TCAATTGCTAC AATGAAAAAG AAGGGTGAGA AGAGATGTCT GAATCCAGAA 300  
 TGAAGGCCGA TCAAGAAATTT ACTGAAAGCA GTTAGCAAGG AAGTGTCTTA AAGATGCTCT 360  
 TAAATCCAGA GCGTGCAGAA ATGATATGCA TCGTCCAGG GATGGACAC ACAGAGGCTG 420  
 CCGTCCCAT CAGTTCCTTA CATGAGTAT ATGTCAAGCC ATAAATGTTC TTAGTTTGCA 480  
 GTTACACTAA AAGGTGACCA ATGATGGTCA CCAAAATCAG TGCTACTACT CCGTATGGAA 540  
 GGTAAATGTT CATCATCTTA AGCTATTCCG TAATAACTCT ACCGTGGCAC TATAATGTAA 600  
 GCTCTACTGA GGTGCTATGT TCTTAGTGGA TGCTCTGACC CTGCTCTAAA TATTTCCTCT 660  
 ACCTCTTCCA TCTTCCAAGG GTACTAAGGA ATCTTCTGCT TTGGGGTTTT ATCAGAATTC 720  
 TGAGAATCTC AATAATACTA AAGGTATCCA ATCAAAATCT CTTTTAAAG AATGCTCTTT 780  
 ACTTCAATGA CTTCACCTGC CATCTTCCCA AGGGGCCCAA ATTCCTTTCAG TGGGTACTGA 840  
 CATACAAATC CAACACATA CAGGAAGGTA GAAATATCTG AAAATGTATG TGTAAAGTAT 900  
 CTATTTTAAT GAAAGACTGT ACAAAATATA AGTCTTAGAT GTATATATTT CCTATATTGT 960  
 TTTCAGTGTG CATGGAAATA CATGTAAATTA AGTACTATGT ATCAATGAGT AACAGAGAAA 1020  
 TTTTAAAAAT ACAGATAGAT ATATCTCTGT CATGTATCAT AAGATAAATG TGCTGAATGG 1080  
 TTTTCAAAAT AAAATGAGGT ACTCTCTCTG AATATTAAAG

Seq ID NO: 104 Protein sequence:  
 Protein Accession #: NP\_001556.1

30 1 11 21 31 41 51  
 MNQTALICC LIFLTLSSQI GVPLSRTVRC TCISISNPV NPRSLEKLI IPASQCPRV 60  
 EIIATMKKKG EKRCLNPESK AIKNLKAVS KIMSKRSP

Seq ID NO: 105 DNA sequence  
 Nucleic Acid Accession #: NM\_015068.1  
 Coding sequence: 1170-2243 (underlined sequences correspond to start and stop codons)

40 1 11 21 31 41 51  
 GTAAACAACG TCACCTGGG TCGGACTGC CCACTTCTC CTCTCTCCC TCCTCCCAAC 60  
 AACACAACAA ACAACAACCT CAAGCACAC GGCCTAAGA GTGCGTGTGT CCCCACATG 120  
 ACCGAAACGAA GAAGGGGAAGA GCTCTGTGA GAGATCAACA ACTTAAGAGA GAAGGTGATG 180  
 AAGCAGTCCG AGGAGAACA CAACTCTGAG AGCCAGGTGC AGAAGCTCAC AGAGAGGAAC 240  
 ACCACCTCTC GAGAGCAAG GAAAGCCACC CCGTAGATG AGGATGATGA CATGCACTC 300  
 CCGCTGTCTG CAGCACTGCT TGCGCCACCC CCGCTCAATG AGGAAGAGTG CCGAGAAGAC 360  
 CTCCACAGAGA AGTTGATGTG CAACCGAGAC ATGCTGGCTC CTTCATGGC CCAATGCCAG 420  
 ATCTTCATGG AAAAGAGCAC CAGGGATTTC TCAATTGATC GTGTCCGTGT CTCTCTCTG 480  
 AAGACCATGA TGACCGGCCG TGCTGCGCGT TGCGCTCAG CAAAGCTGGA GCGCTGCCAC 540  
 TACTGTATGC ACAAC TACC AGCTTTGATG ATGGAAATGA AGCAATCTTT TGAAGACCTT 600  
 CAGAGGGAG AGCAAGATCA AGCAAGATC AGAGCGCTGC GCCAAGGACT GGGGTCTGTC 660  
 GTGACTACT CCAATGCTTT CCAAGATGATT GCCACGAGCC TGGATTGGA CAGAGCTCGC 720  
 CTGATTGACC AGTACACAGA GGGCTCTAGC GACCAACTTC AGGAGGAGCT CTCCACCTC 780  
 GAGGTCCGCA AGTCCGCTGT TCGTCTGATT GGGCAGTGCA TTCACATTGA GAGAAGGCTG 840  
 CCGAGGCTG CTGCACTGCT CAAGGCCAGC TGCCACCC GGGGCTCGGT GTTCTCTCATC 900  
 ATTCCAGCC AGCAAGATCA AGCAAGATC GAGCGGTGAG GAGTGTGCCG CATCTCCCTG 960  
 AGCGAGGAG AAAGAAGAA AGCGCAAGAA CTGAACCTGT GGCCTACTCT TGGAAACAGGA 1020  
 GGTCACTACG CTGACAAATT TCGTGCCAAG GCGTCAAGT CTTCGCGCGC GGGAACTTC 1080  
 CCGGCGCCCG TGTAGAGGGA CCTTCAGGGA CCGGGCCAGA AATAATAAG TCCCCACAAG 1140  
 ATGATGCTCT ATCTCCACAC TTGCAAGTGA TGCTCCAGAT TCACTCTCC GGCAGAGACA 1200  
 CCGTGTGTGT CCGAGGCTAT ATCGATTGTG GTGCTCTGAG CAACCTATT GATCAAGAA 1260  
 ATTTTGCTCA AATATGAATT CTTTAAAGA TGAAGATCT GCGCAATCTT GTTAGAGACA 1320  
 TTGATGGGCG CCGCAATGCA TCGGGCCAG TTGTCACGA AACCTACGAC CTGATAGTTG 1380  
 ACCCTGGGGA TCACCGAGAG GTGCTGTGAT TTGATTGAC TCACTCTCCA TCTTCCCTG 1440  
 TGCTCTAGG GGTTCGCTG CTGAGCACAC ATGATCCCAA TATCACATGG AGCACTGGAT 1500  
 GTATCTCTT TGAATTCGAA TACTGCGCT ACCACTCCG GATGTATGTT CCAATACCA 1560  
 CATGCGCTGC AGCAGAGCA CCAACAACCG CACTCTATA TCCATATGAT GATACAGAG 1620  
 TTACCAACC AGTATGAT TACTATGTTC AGAATGTGTA CACTCCAGTA GATGAGCAG 1680  
 TCTACCCAGA TCACCGCTGT GTTGACCTC ACATAGAAAT GATACCTGGA GCACACAGTA 1740  
 TCCCGATGG ACAATGTGAT TCACTGTCCG AACCTGAAAT GGCAGCTCT CGAGATTITG 1800  
 TGGCAAGAAA TGTAAAGAT GGGCTAATTA CTCACAGAT TGCACTTAAT GAGGCCAAG 1860  
 TCTCCAGGT GAGAGAGGGG TGGAAACTGC AAGTTCTCTA TGAATTCGGA GCTTCAACA 1920  
 ATTTTACTAT CCAAGATTCAG TATCTCGCGT TATCTATTC AATTTAGAA GACCAAGCAC 1980  
 ACCGGGCAAC GTACACTGAA TTGTAACCTC AAATACCTGG ATACCAAAACA TACCCTACAT 2040  
 ATGCGCGTA CCGAGCTTAC CCAATAGGAT TCGCTGGTA CCGAGTGGGA CGAGACGGAC 2100  
 AAGGAAGATC ACTATATGTA CTTGATGAAA TCCACACTGG TACCOCGAC 2160  
 CTCGGTACC ACAGTACCG CCGCACAGC CCGCGCTCC CAGTACAGA CGGCGCGC 2220  
 CTCATCTTA CAGTACCGG TAAATACCTG TCAATCTCT CAGGATCTCT GCGCTCAAAA 2280

TTATCTCTG TTACAGCTCT CAATCACTGA CTGTGCTCT AATTTTAGCTACTGTATCT 2340  
 TCAGCCACC TGAGGACACAT CTTCTCTGAA ACGGCTATGG AAGGTATGG CACATCTGGA 2400  
 GTGGCACACA TCTTAAGCA CCAAAAGACC TTCAACATTT TCTGAGAGCA CAGAGATATT 2460  
 TGCCAATAAA TGATCTCTCA TTTTCCACC TTGACTGCCA ACTCAACTAA AATAATTAAT 2520  
 5 AGTTTACTT TCAGGCGABT CCTGGGAAGTC TGGGTTTAT CTGCGAAAC CTCCATCAC 2580  
 ATCTAAATTA TAGCTGSCCA AATTGCTGT TTAACATTTA CAGAGAAGCT GATACAAAGC 2640  
 CAGGAATTC TGATCTCTT ATGAGGGGG AGAGAGAGG GAGGAGAGCA TGACTTTCT 2700  
 TGGGTTTG CTACCTCTT TTAAATCAC TGGAGGACTG AGGCTTAT AAGGAAGCCA 2760  
 10 AAATTAATGG TSCAGTGTG AAGGCTTCC GTGATCTCTT CGCTGCACC TTAGAAGCTT 2820  
 CACCGCTCTT AAATCCAT TTCAATGGTTC TGTTAATTC CAAGGAGCAG CAATCGACT 2880  
 GGTCTGCCA GGAGCAGGAA AATCCCTGT GACATGAAC ATCTCAGGCC TGAAAGAGAA 2940  
 GTGCTCTCT AGATGGACTC TTGCATGTTA AGACTATTT TGCATCAT GTTGCAATC 3000  
 ACATGACC AATGACTCG GTTTGACAC AACACCTTAC CATCATCATG CCATGATGGC 3060  
 15 TTCCACAAAG CATTAACCTT GGTAAACGAG GATTACTGTT GGCTCCAGCG TTGTATAGAT 3120  
 TFCATGAAT GTGACACCT CTCAATCAC TTTGAGGGCT AAAGAGTAGC ACATCAAAAG 3180  
 GACCTCAAAA TCCCATACCC AACTCTTAAG AGATTGTGTC TGGTACTTCA GAAAGAAATT 3240  
 TCATGAGTGT TCTTAATGCG CTGGAAAGC ACCAGCTGAC GTTTTGGAG AATCTATCCA 3300  
 TGCTGTGCG TCCATATGCA TGTGGGCATT TCAATCTCAG TCCCTCATT AGACTGTAGC 3360  
 20 ATTAGGATGT GTGGAGAGAG GAGAAATGAT TTAGCACCCA GATTTCACAT CTAATGCCGT 3420  
 GAAAGGGGAC ATCTTTGAG AAGAGGAATT AGGGCTGTGG ACAGTGTCT GAGGATGTGG 3480  
 ACTTCCTTAG TGAGTCCATC ATTACTTGAT GTTAACCATC TCAAAAGGAT CAGATCCAC 3540  
 GTAATGAAA AGTTCCTCT AGAGATATGA GCTGTATGTA AGTTCACCAT GGTATGAAAG 3600  
 CCTCAGAAAG CAATCAACG AATGAGCAAC AAGGAGCAAC ACAAGATTTT CTCTAACCC 3660  
 25 AGTGACACT CTGATGTCC CTGGAAGCTT TGTGCTAAC TGGGACTGCG TGACTCTTCT 3720  
 TAGCCTGGTC CTGCTTACT ACCTGAACCT GTTTATCTA ACCTCTCTT TCTGTCTTAA 3780  
 TTCTTGCTA CTGGCATGA CCTGCTGCA GGATTGTGT CATTTCTCG CCGTGTCTT 3840  
 GAGACTCCAT TTGCTGCCA CACACAGAGA TGAAGAGGCG AGGCTTAA TGTCCAAAGCA 3900  
 CAGTTTGAAG AGTAGAAAC AACATGGTGT ATATCTCAAA TTGCTGACA TGAAGAGAGC 3960  
 TCTAACCGT ATGATTTTCT TTTCATCAG ATCACTTTC ACATGTTTAT TATCATCCGC 4020  
 30 TCTATTCTT GCACTTTTAA ACACCTAAAA TTTTATGAT AATTITTAGT TGTGTTGAA 4080  
 GTGGTGACTA GCGTTTCAAA AACTTCCATT GAATTACAAA GCACTATCCA GTTCTTATTG 4140  
 TTAACATAAG TAAAAATGAT AAGTAACATA GTGTAATAA TTCTTTACT GTGACTCTT 4200  
 TACAATGCTG TGAATGAGAG GCTCCTCAGA ACTGAGAGCAT TTGTATAATA ATTCATCTG 4260  
 35 TTCACTTCA ATTTATGAT CATATAAAT TCAATATCTA TCAATTTGGC CTTTAAAGC 4320  
 CATATAAAG GATATAAAT TTGAAGAAG AAOCTTAAAT GCGTATTTAA TCCAAAGCAA 4380  
 CTTTCTTTT TCCITCAATG GAATCAGAAA GCTTGTCAAT CACTCATGTG TTTTAGAGTA 4440  
 ATACTTTTA AATGTGTGCA TTTGTGCTTC TGAACATTT TGAAGAGTCA CTCTGTITA 4500  
 40 CCTCAAGTA CAATTCATCC TCCATACATY TGAATTCAAG TGTGTTTTG TCAAAATTAC 4560  
 AGTTGTCAAT TGATCTTCA AGTCCAGGGT GCTGAGAAAT GGGCTCTTCT CTATGAGCT 4620  
 GGCATGTGCA CAGGCAACT TCCACCACT CCAAGCAAA GTTCTGAGAA GCTCACCAAC 4680  
 CACAAGAACG ATTAGGAAA ATATGCTGCT GTGGTTTAA CACTCAGAAA GTCCCTGATC 4740  
 CACATTGGC TGTTTACIA AGCTGTGAT TAACTTTTG CGAGTGTGA CTATGCTCTA 4800  
 45 TGTCTATATA TGCTATCTAT AATGTAGAT GTTAAGGATA AGTAATTCTA AATTATTAAT 4860  
 TCTATAGTT TGAATTTTGG TTAAGTTTGC TTTCATCTCA TGTATTAT TGTGTGTTAA 4920  
 TCAATTTAT GTTATTTGGA TCTTTAAAT TTTTGTGCG ATTTCACAC AAAAAAGCT 4980  
 50 TATTTCATAA GAAAGGAAA AAATCAATGG AATTGTATAT CTAAGAAGT TAAAGAGGA 5040  
 GCAAAATAAA AAACATAAAG GAGATAGATG AATTAGTAAG CAAATCAGTA GTGAGTITT 5100  
 TCAAACTGGC AAAATTAAT AATTGACTTT TAGCCCAATT TTACATGTT AATAAATTA 5160  
 55 AGAAGGAGCA AGATTAAGA GTCTCCATTG ATAGGCAAGC CTAGAGAGAA CTAGCTAAAT 5220  
 TTTATCTCT AGGACTCTA ACACAGAAA GTTACATAC ATTTATGAG GTTCAATTA 5280  
 GTTTGGACAG TGAGTATTT GTCTTATGTC AAAAAAGAG AATTAGTCT ATCAAAATCT 5340  
 GAAATATAC AGTGAACITG CAGGTGCACA AAATAAGAGG GGCACATCTA TATGGTGACG 5400  
 TCTGGAATTC TGTTTAATT TGTAGTACG TCTTGACATT CTGAATTGAT CCAATTGTCA 5460  
 60 TCCACCACAG ACATCTACA CAGATAGAC ACAGTTCCAA GATTGACAC AGAGAACAAC 5520  
 CTGCTGAGAA GACTCTGGCA GAAATGGAGA GCCCTGGGGG AACCATCTTA CATTTCAT 5580  
 TAAAGAGAGA ATCCACATCT GATGACAGCT AAGGTCTTT GTTGTGTTAG ATTGTAGAT 5640  
 65 GGTATGAAAT TGGTCTGTGG AAAATTCGAT TGCCTTATT TCTTTGTGA ATCAAGATTA 5700  
 AGTAATAGGG GATATATAAT CATAGCATT TTAGGGTGGG AGGGACTATT AAGTAATTT 5760  
 AAGTGGTGGG GGTATTTAG AATGTTAGAA TAATATTTG TATTAGATAT CACTAAGT 5820  
 60 GGCATCTGCT ACTTACTTGT AACCTTTAC CCTATAATTG CTATCCTTAA AGATTCAAA 5880  
 TAACTCTGA GGGACTCTGA GCGAGACCA CTATTATGTA GGGATATGTA GATGATAA 5940  
 AACCCCACTG GGAGAAAGT CAAAGGTGAT TAGATTAATA ATTTAAGA GAGTAGTGA 6000  
 CCTCTGATAA ATTACTCTTA GAATGAACIT GTCATAGATG GATGGTAAAT TTTACTGAAA 6060  
 75 GTTATAAAG TGATAAATA AAAOCCCTGC TTTTACCCT GTACAGGCC CTCTCTCTAC 6120  
 CACTGACACC CATTGCCCT ACCOCTCTCT CTAATCTTAT TCGTGTATC TCTTCATCT 6180  
 ATATTCTCT CTATTGCTA ATATTGCAAT GCTUTTACAA TAAAAATTC AATAAGATT 6240  
 AGTGGTTAAG TGC

Seq ID NO: 106 Protein sequence:  
 Protein Accession #: NP\_055883.1

1 11 21 31 41 51  
 | | | | |  
 MTERRRDEL S REHNLRK E V MKQSENNNL QSQVKLT E NTLRQVEP TPDEDDDDIE 60  
 LRGA AAAAAA PPIHEECP E DLPEKGD GNF DMLPFAQC QIFMEKSTR D FSDVRVRVC F 120  
 VTSMTGTG RAA RWASAKL E RHYLMHNYA F MMEMKHYP D FQRREVAKK R IRLRQGM S 180  
 VIDYSNAP Q IADQLDWN E ALIDYQRE G L SDHQELSL L EVAKSLSL IGQCHERR R 240

LARAAAARKP RSPPRALVLP HIAHHQVDP TEPVGGARMR LTQBEKHRRR KLINI.CLYGT 300  
GGHYADNCPA KASKSSPAGN SPAL

Seq ID NO: 107 DNA sequence

Nucleic Acid Accession #: NM\_003679.1

Coding sequence: 47-1507 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
10 GCACGACGA GAAGCAACA TAATTGTGAA AATACTCTA GCAGTTATGG ATCTCATGTG 60  
CATCTAAAGG AAAAAAGTAG CTGTCAATGG TGGTGGCTGT GTTGGCTCAT TACAAGCATG 120  
CTTTCTTGCA AAGAGGAAT? TCCAGATTTGA TGTATATGAA GCTAGGGGAA ATACTCGAGT 180  
GGCTACCTTC ACACGTGGAA GAAGCATTTAA CTTAGGCCCT TCTCATAGAG GACGACAAGC 240  
15 CTTGAAGAGCT GTTGGCTGGG AAGATCAGAT TGTATGCCAA GGTATTCOCC TGAGAGCAAG 300  
AATGATCCAC TCTCTTTTCAG GAAAAAAGTTC TGCATATGCC TATGGGACAA AGTCTTCAGTA 360  
TATCTTTCT GTAGCAGAGG AAAATCTAA CAAGGATCTA TTGACTCTGT CTGAGAAATA 420  
CCCCAATGTT AAAATGCAT TTAACCACAG GCTGTGTGAA TGTAAATCCAG AGGAAGGAAT 480  
GATCAAGTGT CTTGGATCTG ACAAAGTTCC CAAAGAATC ACTTGTGACC TCAATTGAGG 540  
ATGTGATGGA GCTTATCTAA CTGTCAAGTC TCACCTGATG AAGAAACCTC GCTTTGATTA 600  
20 CAGTCAGCAG TACATTTCTC ATGGGATACAT GGAATGTGACT ATTCACCTTA AGAACAGAGA 660  
TATAGCCATG GAATCTAATT ATCTCATAT TTTGGCTAGA AATACCTTTA TGAATGATGC 720  
ACTTCTTAAC ATGAACAAT? CATTCACATG TACTTTGTTC ATGCCCTTTG AAGAGTTTGA 780  
AAAACTCTTA ACCAGTAATG AITGGGTAGA TTTCTCCAG AATACTTTC CGATGCCAT 840  
CCCTCTAATT GGAGAGAAAC TCTAGTGTCA AGATTCTCT CTGTGTGCTG CCGAGGCCAT 900  
25 GATATCTGTA AAGTGTCTCT CATTTACAT TTAATCTCAC TGTGTGCTG TGGGAGATGC 960  
AGCTCATCT ATAGTGGCTG TTTTGGGCA AGGAATGAA? GCGGGCTTTG AAGAATGCTT 1020  
GGTATTTAT GAGTATGAG ATAAATCTAG TAAAGACCT? AGTTTGTGTC TTCTCTGTGT 1080  
CTCAAGATTG AGAATCCAG ATGATCACGC GAITTCAGAC CTATCCATGT ACAATTACAT 1140  
AGAGATGCGA GCACATGTCA ACTCAAGCTG GTTCAATTT? CAGAAGAACA TGGAGAGATT 1200  
30 TTTCTATGCG ATTATGGCAT GGAOCTTAT CCCCCTCAT ACAATGGTCA CTTTTCAG 1260  
AATAAGATAC CATGAGGGCT TGCAGCTGTG GCAATGGCAA AAAAGGGTGA TAAACAAGG 1320  
ACTCTTTCT TTGGGATCA CACATACCAT CAGCATGCC TACTCTCTA TACATCATAT 1380  
ATCACACGA TCTTCTCTCT GCTTGAGGAG ACCATGGAAC TGGATAGCTC ACTTCCGGAA 1440  
TACAACATGT TCCCCGCAA AGGCCGTGGA CTCCCTAGAA CAAATTTCCA ATCTCATTAG 1500  
35 CAGGTGTAG AAAGGTTTTG TGTATGCAA TGCATGATT CTCGTGACC AAAATTAAGC 1560  
ATGAAAAAAA TTTTTCAT? GCCATATTTG ATTCATAGT GGAAGATAGT GTTGCTGTCTA 1620  
TATAAATCT GATGTAGAG TATCTCTGTGA TGTATAGTC AATTAATCTGT TGGTGGGTGC 1680  
ATTTTAAAG ATGCAAAATG CAGCTTCTCT ACATATACGA CACTCAGATG GTTGATCTCT 1740  
TACTATAAAA GTGCAATGAC TAAGATCTCT CACTTCTGT AAGTAAGGC CTAATGATCC 1800  
40 AACGGAAGA CAGTAATCAT GCTTTTCTT TAAAAGACAC AATAGGACTC GCAACAGCAT 1860  
TGACTCAACA CTAGGAGCTA AAAATCACA CTAACTAGTC ATGTTAAGTC CACTTTTACT 1920  
TAGGTGAATG GAATCTTACT AACCACAGGC CTAGACTTCA CTAGATAAAA CCAGAAATGT? 1980  
AAATAAGGAA TCGAGGGAG TCCAGAGAC TTACAAATG AACTCATTT? ATTTTCCAC 2040  
CTCAAAATAT AAGTATTAIC ATCTATCTGT TATCTGTCTA TCTATCTAT ATCTATCTAT 2100  
45 CTATCTATCA TCTATCTATC TATCTATCTA TCTATCTAT TATCTATCTA TCTATCTAT 2160  
TCTATTTAT TATGATTTTA GAGATCAGGT CTCACCTGTG TGACCAAGGT GGAATGCAGT 2220  
GTTGATGACT GGGTCTACGT CAACCTCTGC CTCTTGGGCT CAAGAAGCTC TCCACTCTCA 2280  
GCCGCCAA TAAGTGGGCT TACCATGTGA TTTTCTGTA GAGAGCCGGT CTTCCTATGC 2340  
TGCCAGGCCG AGTCTCAAC TCTTGGCTC ATGTGATCTG CCAACCTCAG TCTTCCCAAG 2400  
50 TACAGGGATT AAGGTTTGTG GCCACCGCTG CCAGCCGAGA GTTACCCTT AAGATTAAGA 2460  
AAAAGGCTAT TAATATCAT CTAAGTGAAG QACAGGAGAG GGTTTTATTC ATAAATTAAG 2520  
TGTCTACATG TGCCGAAGTG GAAAGGAAC AACGGGGAGAC AACTTTTATA GAAATACAAA 2580  
GCCATTAAT TATTCAATT TGAGCTCTCA GAAGCAATTT ACTAATATT TCTGTGACTA 2640  
CATATCTGCG CAGACACAG CATACTGCTG ATTTTATAA GCAATCTTAG TGAATATGTT 2700  
55 TCTTTGGTCT ATCTTCTT? AACAGGCTGC TGAGTCACTC AGAAATCT? CAACATGAT 2760  
TAATTATGAA GATGAACAC TAGAGTCTAT TAAGAAATAA AAATTGGGCA ATAAATATAA 2820  
ATGATTCAGT GTTCTTTTC TATATTTGTA ATGAAAGCT? TGAATCTAA TAAITCAATG 2880  
TATGTTTGTG GGGAAAGAAA AAATAATTT? TCTCTTACT CACTTTAGGT? TCTTGGGCTG 2940  
GCGGCCAT? AACA AAGAA AGATGAGAA GAGAAACAA ACATAAAT? TATTATGGGG 3000  
60 TATATGTAAT ATATATGTTG GAATATCAG GBAATGAGCA AATCTCAAG AGCTGGGCTC 3060  
TGAAGATCC CTGGCTTATA TAGCATGAC AAGAAGACAT AATTTTTAG AGAAACACA 3120  
AAACAAGAAA AAAGAGCTTT GAGTCTGTAG GGGCAGCAAT TTTGGGGGAG CAAATATATG 3180  
GGAGTTTGGC TGTGATGCT CTCGTGGTGT GGTCTCAAG CTGACAGAGA TCCAAGTTG 3240  
TCTGAAAC TGGCTTCTG CTATCTGAC ATATAAAGC GTTCTTTGT? CCAACAGAG 3300  
GATTTCTTT? TCAATCTAGA ATTATCTCT TGAATCTAT ATCAGATTA GAGACATGAG 3360  
65 CTGAATAGAG TCCAACAGTA CAAAAAAAT? TCAATGATGT? CTAGCTACT? CACACATGTT 3420  
TAGCGCAGAG TTAATTTTAC AGTAAGGAT? TTTGAGAAA AATGCATTAC GTGTTTGGAG 3480  
AATAGAGTA ATTTAAAAA TATATTTGAA ATGAAAATCT? CCAACACAT? AGAAGATGAT 3540  
GATTTTATG GGCATGCTG TGCCCAAGG GTTTTTCTA TATGTATAAG CACCGGTGAG 3600  
70 ATTAAGAA TGTTTTGTG TCAACCTCT TGTGAGGCT? CAGAGACATA TGGGGAATC 3660  
GGATTTCTGT CTGACACCA AGTGTCTGTT GATATTTAC ATGAACCGGT? ATTCACAGCA 3720  
CTTCTTATGA TGAAGATCCA AAGTGGCAT CCAATTTAAG GCGCCATCT? TGTGTGCCAT 3780  
TCTCTATCTC TACAAGGAG CAACTGGAT TATGATCAT TTGAGCCAT? TGGTTTGTG 3840  
CGTGTGCTCA ACTGAGAGT ATTCACTACT GGTGATGAT AAATATATG AAGAGAGGCT 3900  
75 GAAAGTCACT TGTGTTCTCC ATCTCTCTC GAGATGACA TGGGCTCTGAT 3960  
CTGACATTTA CTTCACGCTG CTGATAGTCT TTAGCAGGC CTCGCGACC TCGAGCTGT 4020  
GAGGACAGA AGCTGC AAAA GGGATCTTG AACTGGGCA GAGAAAAAT AAGATGGAAT 4080



ATTAAGTAAA AGTTGGGCAC TAATCTGGAT TAACATTGGA GGAATCACT TGAGCTGATT 4140  
 TAAGTTGTTT TTTGTTGTT AGCAGGTGTG GATGTGGGGT TATGTGGTCA TGCTCAGATC 4200  
 TAACATAATC ACCCGACAGC TTATGTCTTT TTATTCATTC TAAATCTTAT TAACCGGAAT 4260  
 ATGTAGGAAC ATTCAATAC CTGTGAATCC TCCAAAGCTTC AATCTGCACA CACTTTGTAT 4320  
 GAGGCGAGGT ACAACTATTA AGAGATTTTG AACATAAGT TAGTCCACAA ATATTAGTGT 4380  
 GGCACTACT AGGTGACAGC CACTCTGTCTA TAATTAGAGA CTTTTACTA TAAGATCAA 4440  
 AAACAGATAA GCTCTTCTCT GGCAGAGTTT ACAGCTGTGT GTACTTGCTA ATGTCTTCTT 4500  
 AATTAGTGGA AGAATTTT TTTCTATCG AATATCTAA TCAGTTGGGG AAAAAATATC 4560  
 CATAGCAGAC AGCACTAATG TCATCAACAA ACATTTGTTCT TCTCGTGTG CTGGGTACAA 4620  
 CATOGAATAA TATTCTTGG CTTCTTCTCC GCTCTCTCC TCTGTGTTC CTCTACAAA 4680  
 GAACCTGGGA GGCCACAGCC TAAAGATCAT AATATCAAA TGAAGGAGAC CTAGATCTCT 4740  
 AATAGCTCTC ATAGCACTA TCCATCTCC TCCATCTCC ACATTTATG ACTGAGATCT 4800  
 GACCTGAAAT GAGCAATAAA CTCTGTATTA ATTCACTGAA ATGTTGGGT TGCTTGATAT 4860  
 AGTAGTGGT CCATCATGAC CAGTAAACAA TAAATCAAAA GTTAATGTAA TTGTTATCCC 4920  
 ATATTATGA GCGAAATAA TTGGAATAT ATGGACTTTC TCAGATTAGG AAATACCAAT 4980  
 TAAAAATATA ATAAATAGCT

Seq ID NO: 108 Protein sequence:  
 Protein Accession #: NP\_02670.1

1 11 21 31 41 51  
 | | | | |  
 MDSSVQKXK VAVIGGGLV SLQACFLAKR NFQIDVYEAR EDTRVATFTR GRNINLAISH 60  
 RGRALKAVG LEDVQSQGI PMRARMHHSI SXKKSAPYQ TKSQYLSVS RENLNKDLIL 120  
 25 RGRYPNVKM HFNHRLKCN PEEGMITVLG SDRGPKDVTIC DLVCGDGA YSTVRSHLMKK 180  
 PRFDYSQVY PHOYMELTIP PKNGDYAMEP NYLHVPRNT FMIALPNMN KSFCTLEMP 240  
 FEFEKLLTS NOVDFPQKY FPDAPLJGB KLLVQDFPLI PAQPMISVCK SFEHFKSHCV 300  
 LLDIAIAIV AGTCTCTCTC FBDCLVTDEL MCKFSRHSI CLPFSRLSI PDDHABSLG 360  
 MYNYEMRAH VNSSVFWQK NMERPLHADM PSTHPLYTM VTPSRIRYHE AVORWHSYQKK 420  
 30 VINKGLFPLG SLIAISSTYL LHYMSRPSF LCLRLPWNWI AHFRNTCFP AKAVDSHLEQI 480  
 SNLISR

Seq ID NO: 109 DNA sequence  
 Nucleic Acid Accession #: NM\_006115.1  
 Coding sequence: 236-1765 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 | | | | |  
 GCTTCAGGTT ACAGCTCCCC GCGAGCCAGA AGCCCGGGCTT GCAGCCCTTC AGCACCCTGC 60  
 CGGACACACC CACCGCTTCC CCAGGCGTGA CTTGTACA CA GCACTTCGC GGTGTGGTGA 120  
 40 ACCTCTGAG GAAAAACAT TTGATTATT ACTCTCAGAC GTGCGTGGCA ACAAGTGACT 180  
 GAGACCTAGA AATCCAAAGC TTGAGGCTCC TGAGGCCAGC CTAAGTGCTT TCAAAATGGA 240  
 CCAAGAGCGT TTGTGGGTTT CCATTCAGAG CCGATACACT AGCATAGTGT TTGTGACAAAG 300  
 CCCAGCGAGA CTGTGTGAGC TTGAGGGA GAAGCTTCTGT AAGBATTGAG CCGTGCCAT 360  
 45 TCGCGCCGTG GAGTTGCTCC CCAGGGAGCT CTTCGCCCA CTCTCAIGG CAGCGTTGA 420  
 CCGGAGACAC AGCCAGACCC TGAAGCAAT GGTGCAAGGC TGCCCTCTCA CTTGCTCC 480  
 TCTGGGAGTG CTGATGAAGG GACAACATCT TCACTGGAG ACCTTCAAG CTGTGCTTGA 540  
 TGGACTTGAT GTGCTCTTGG CCCAGGAGGT TCGCCCGCAG AGGTGGAACAT TCAAGTGCT 600  
 GGAATTACGG AAGAACATCT ATCAGGACTT CTGAGCTGTA TGCTCTGGA ACAAGGCCAC 660  
 50 TCTGTACTCA TTTCCAGAGC CAGAAAGCCT TCAAGCCATC ACAAGAGAGC GAAAGTGA 720  
 TGTTTGTGAC ACAGAGGACG AGCAGCCCTT CATTCACGTA GAGGTGCTCG TAGACATGTT 780  
 CCTCAAGGAA GGTGCTGTG ATGAATTGTT CTCTCACTC ATTGAGAAAG TGAAGCGAAA 840  
 GAAAAATGTA CTAAGCCTGT GCTGTAGAAA GCTGAAAGATT TTGCAATCG CCATCAGGA 900  
 55 TATCAAGATG ATCTTGAAAA TGGTGAGCT GGACTCTATT GAAGATTGG AAGTGACTTG 960  
 TACTGTGAG CTAACCACT TGGGAAATT TTCTCTTAC CTGGCCGAGA TGAATTAATCT 1020  
 GGTGAGACTC CTCTCTCCC ACATCCAGCG ATCTTCTACT ATTTCGCCG AGAAGGAAGA 1080  
 GCGATATATC GGGCAATCCA CTTCTCAGTT CTTCAAGTGT CAGTGCTCG AGGCTCTCTA 1140  
 TGTGAGACTCT TTAATTTCC TTAGAGGCGC CTTGGATCAG TGCTCAGGC AGGTGATGAA 1200  
 60 CCGCTTGAAA ACDCTTCAAA TAACATACT CCGGCTTCC GAAGGGGATG TGATGCATCT 1260  
 TGTCCAGAGT CCACAGCTCA GTACATTAAG TGTCTCAGT CTAAGTGCGG TCATCTGAC 1320  
 CCAATFAGT CCGAGGCCCC TCAAGCTCT CTTGAGAGGA GCTCTTGCCA CCGTCCAGGA 1380  
 CTTGCTCTTT GATGAGTGT GGAATCAOGA TGATCAGCTC CTTCGCCCTC TGCTTCCCT 1440  
 GAGCCACTGC TCCAGCTTA CAACCTTAAG CTCTCAAGG AATTCATCT CCATATCTGC 1500  
 65 TCTCGAGAT CTTCTGCAAG ACCCTATCG GCTGAGCAAT CTGACCCAG TGTGTATCAT 1560  
 TGTCCXCTCT GAGAGTTTATG AAGACATCCA TGGTACCTC CACTGAGAGA GGTCTGCTTA 1620  
 TCTGCAATCC AGCTCAGG AGTTGCTGTG TGAATTTGGG CCGGCCAGCA TGGTCTGCTG 1680  
 TAGTGCACAC CCGTCTCTCT ACTGTGGGGA CAGAACCCTT TATGACCGG AGGCCACTCT 1740  
 GTGCCCTGT TTACTGCTTA ACTAGCTGGG TGCAATATCT AATAGCTTCA TTCTGCATAC 1800  
 70 TGTGACACTA AAGCCAGGAT GTGATCATCT CTTGAGCAAA CAAAGCAGCC ACAAGTTGAG 1860  
 ACAATGTTCT AGTGTGAGTG AGGAAACAT GTTCAGTGAG GAAAAACAT TCAGACCAAT 1920  
 TTGATGTAAG GAAATAAGG GGAAGTTGGG GTATGAGGCA TGTGACTGT AGGAGTTAT 1980  
 GTGATCTTTG GGAAGACTA TCTTATAGAG TTAGAATAG AATCTGAAT TCTAAGGGA 2040  
 GATTCCTGGT TGGAGATGAC ATGTAGGAGT TAATCCCTGT GTAGACTGTT GTAAGAGAAC 2100  
 75 TGTGAAAAA AAGAGGAAGC AATGTGAAGC AAAAAAAAAA AAAAAAAAAA

Seq ID NO: 110 Protein sequence:  
 Protein Accession #: NP\_006106.1

1 11 21 31 41 51  
 5 MERRRLWGSIS QSRYSIMSVW TSPRLVELA GQSLKDEAL AIAALELLPR ELPPFLFMAA 60  
 PDRHSISQTLK AMVQAWPTFC LPLGLVMKGG HLHLETFKAV LDGLDVLIAQ EYRPRKWKIQ 120  
 VLDAKRNQGR DFWFTWGGGR ASLVSPPPEA AQRMTKKRK VQGLSTEAQ FPPFVLYD 180  
 LFLKEGACDR LFSYLIEKYK RKSNVRLCC KLELFAMPM QDRKMLKMY QLDSDIELV 240  
 TCTWKLPTLA KFSYLYGQMI NURRILLSH HASSYISPEK EBQYIAQFIS QFSLQCLQA 300  
 10 LYVDSLFFLR GRIDLQLLRHV MNPLFTLSIT NCRLSBGDMV HLSQSPSVSQ LSVLSLGSVM 360  
 LTDVSPEFLQ ALLBRASATL QDLVDFBQGI TDDQLLALLP SLHSCSQLTIT LSPYGNISII 420  
 SALQSLQLHL IOLSNLTHLV YVPFLESYED IHTOLHLERL AYLHARLREL LCELGRFSMV 480  
 WLSANPCTHC GDRFTPYDPE ILCPCFMFM

15 Seq ID NO: 111 DNA sequence  
 Nucleic Acid Accession #: NM\_003815  
 Coding sequence: 8-2452 (underlined sequences correspond to start and stop codons)

20 1 11 21 31 41 51  
 CGCTGCCATG CGGCTGGGCG TGCTCTGGGC CTGGGGGCTC CTGGGCGGGG GCAGCCCTCT 60  
 GCGTCTGCG CCGCTGCCCA ATATGCTGCG CACAGGAG GAGCAGGAG AGTCAGAGAA 120  
 GGCCCGGAGG GAGGCGCTTG AGGCCCAAGT CCTTCAGGAG GATCTCCCAA TTAGCTCCAA 180  
 AAAGGTGCTT CAGACCAGTC TGCTTGAGGC CTTGAGGATC AAGTTGGAGC TGGACGGTGA 240  
 CAGTCATATC CTGGAGCTGC TACAGAAATG GGAGTTGGTC CCAGGCCGCG CAACCTGGT 300  
 25 GTGTGACCAAC CCGATGGGCA CTCGGGTGGT CAGTGAGGGA CACACTTGG AGAAGTGTCT 360  
 CTACAGAGGA AGAGTGGGGG GATATGAGG CTCTGGGTGG TCATCTTGCA CTCGTGTCTG 420  
 GCTCAGAGG GTTGGGCT TCACCCAGA GAGAAAGCTAT ACCTGGAGC AGGGGCTTGG 480  
 GGACCTTCAG GTGCTTCCA TTATTTCGGC AATCCAGAT CTCACCTCG CAGGCCACAC 540  
 30 CTGTGCCCTG AGCTGGGGG AATCTGTACA CACTCAGAGC CCACCAAGAG ACCCCCTGGG 600  
 ACAGGCGCCAC ATTCCCGGA GGGCGGATGT GGTAAACAGAG ACCAAGACTG TGGAGTTGGT 660  
 GATTGTGGCT GATCACTCGG AGGCCAGAA ATACCGGGAG TCCAGGAGCC TGTAAACCG 720  
 CACACTGAAA GTGGGCTCT TCTGGACAC ATTCTCCGG CCGCTGAATG TACGATGGC 780  
 35 ATAGTGGGCG CTGGAGGCTT GGACCCAGCG TGACTGGTG GAGATCAGCC CAAACGCCAG 840  
 TGTCAACCCTC GAAAACTTCT TCACCTGGCG CAGGGCAGAT TTGCTGCTC GATTGCCCAA 900  
 TGACAGTGCC CAGCTGGTGA CTGGTACTTC ATTCTTGGG CCTACGGTGG GCATGGCCAT 960  
 TCAGAAATCC ATCTGTCTCT CTGACTTCTC AGGAGGTGGT AACATGGAGC ACTCCACCAAG 1020  
 CATCTTGGGA GTGACCTCTT CCATAGCCCA TGAGTTGGGC CACAGCCCTGG CGCTGGACCA 1080  
 TGATTCTGT GGAATAAGCT GCGCTTGTCT AGCTTCAGCC CCAGCCAGAA CTCGTATCAT 1140  
 40 GGAGGCTTCC ACAGACTTCC TACCAGGCTT GAACTTCAGC AACTGCAGCC GACGGGCTCT 1200  
 GGAGAAAGCC CTCTGGATGG GAATGGGAGC CTGCGTCTTC GAAAGCGTGC CTAGCCCTACC 1260  
 CCTATGGCT GCTTCTGGG GAAATATGTT TTGTGGAGCCG GGGAGCAGT GTGACTTGGG 1320  
 CTTCCTGGAT GAGTGGGTGG ATCCCTGCTG TGATTCTTTG ACTGGCCAGC TGAGGCCAGG 1380  
 TGCACAGTGT GCATCTGAGT GACCTGTGTG TCAAAATATG CAGCTGCCCC CGTCTGGCT 1440  
 45 CAGTGTGCT CTAACAGAGG GGAATTGTGA CTTCGCTGAA TTCTGCCAG GAGACAGCTC 1500  
 CCAGTGTCCC CTGATGTGTA GCTTAGGGGA TTGGCGAGCC TTGGCTGGG GGCAGCTGT 1560  
 GTGCA TGAC GGGCGTTGTG CCTCTATGC CCAGCAGTGC CAGTCACTT GGGGACCTGG 1620  
 AGGCCAGGCC GCTGGGCCAC TTGCTCTCA GACAGCTAAT ACTGGGGGAA ATGTCTTTGG 1680  
 GAGCTGTGG CCGAACCCCA GTGCAAGTGA TGTGTCCTGC ACCCTAGAGC ATGCAATTGG 1740  
 50 TGGGACACTC CAGTGCCAGA CAGGTAGAGC CAGCGCTCTG TTGGGTCTGA TCCGGGATCT 1800  
 ACTCTGGGAG ACAATAGAT TGAATGGAG TGAGCTGAAC TGACGCTGGG TGCACTTGA 1860  
 CTGGGCACT GATGTGGCCC AGCCCTCTCT GACTCTGCTT GGCACAGCT GTGGCCCTGG 1920  
 CCTGGTGTGT ATAGACCAT GATGCGAGCG TTGTGGATCT CTGGGGGCCAC AGGAATGTGG 1980  
 AAGCAATGC CATGGACATG GGTGCTGTGA CAGCAACAGG CACTGCTACT GTGAGGAGGG 2040  
 55 CTGGGACCTC CAGTACTCA CACTCAGCT CAAGAGCAAC AGCTCCCTGA CCAAGGCT 2100  
 GCTCTTCAGC CTCTGTGTCT TATTGTTCT GGTGAGCTT GGTGCCGGT ACTGTGACCG 2160  
 TGCCCGCTTG CACCGAGAC TCTGCCAGCT CAAGGGGACC ACTGCCACT ACAGGGGAC 2220  
 CCAATCTGAC CTCTGTAAC GGGCAGGAC CTGCGAGAGG GCGCTGCTGG CAGAGGCCAC 2280  
 60 TAAGTCTGAG GGGCAGAGCA AGCCGCCAC CCAAGGAAO CACTGCTCTG CCGACGCCCA 2340  
 GGGCCCTGG CCGTCCGCG CCGAGGGCT GAGATCCGC CCGATGCTG CCGTGGAGT 2400  
 ACCCTCCAGA CCAAGGCCAC GCGCTCCGAG AGTGTCTCTG CTCTAGCTCT GACTCTTCC 2460  
 GAGGTTCCCG TGCTCCAGC CCGGACTTAG GGTCTCAAGA GCGCGGCTG CCTCTGGAG 2520  
 TCCCTTACCA TGACTGAAG CCGCAGAGAC TGGCGGTGTC TTAAGACTCC GGGCAGCGCG 2580  
 AGCGCTGTCT AAGCAACACT CTGCGAGACT GCGCGGTAG TTGACGCGG GCGTCTGGGA 2640  
 65 GGGGCTGGGG GTTGGACGAG ATTGAGGAAG GTCCCCACAG CCGTGTCTCT CTCAGTTGCA 2700  
 ATAAAGCTGA CATCTTGGGA GCGTAAAAA AAAAAA

Seq ID NO: 112 Protein sequence:  
 Protein Accession #: NP\_003806.2

70 1 11 21 31 41 51  
 MRLALLWALG LLGAGSLPS WFLPNIGGTE EQQARSEKAP REPLEHVLQ DDLFSLKKV 60  
 TLTSLPEFLR KLELDGDH ILLELNREL VPGRPVLVWY PQDTRVSEV GHTLENCQY 120  
 75 GRVKGYSQSW VSICCTSGLR GLVYLTPERS YTLRQPGDIL QGPPIBRIQ DLHLPQHTCA 180  
 LSWRESVHTQ TPPIBFLQGR HIRRLNVDV ETKVIELVWV ADHSEAKRYR DRQHLNRL 240  
 EVALLDITFP RPLNVRVALV GLEAWTQRDL VEISPNPAVT LENFLNRA HLLRPLPHDS 300

AQLVYTSFS GPTVGMQIN SICSPDFSGG VNMDDHSTSL GVASSIAHL GHSLGLDHL 360  
 PGNSCPCYP APAKCTIMEA STDFPLGLNF SNCSRRALEK ALLDGMGSL FERPLSLPM 420  
 AAFCGNMHVE PQECQDCGLF DDCVDPOCDL LTOQLRGAQ CASDGPCCQN CQLRPSGWQC 480  
 RPTRGDCDLP BFCPGDSSQC PFDVSLGDGE PCAGGQAVCM HGRKASYAQ CQSLWGPQAG 540  
 PAAPHCLQTA NTRGNARFSG GRMP8GYSVS CTPRDAICQ LQOCYTRIQP LLGSRDLJW 600  
 BTIDVNYTEL NCSWYHLDLG SDVAGPLLT PGTACRPLV CIDIHQCRVD ILGAQCRSK 660  
 CHGHVUCBN RPTCIBDWA PFDCTTOLKA TSLTGLLL SLVLLVLM LGAGYVYRAR 720  
 LHQRLOQLG PTCQYRAQS GPERGPGP RALLAROTKS QGPAPPPFR KPLPADPQR 780  
 CPSGDLPGG AGIPPLVVPF RFAPPPVTS SLYL

Seq ID NO: 113 DNA sequence

Nucleic Acid Accession #: NM\_020416

Coding sequence: 40-417 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 ATCTCAATACA GGAGTAACTT GGAATCCAT TCTATCACTA TGAAGAAAAG TGGTGTCTTT 69  
 TTCTCTTGG GCATCATCTT CTGCTGTCTG ATTGGAGTGC AAGGAACCCC AGTAGTGAGA 120  
 AAGGGTCGCT GTTCTGCAT CAGCAACCA CAAAGGACTA TCCACTTACA ATCTTGAAA 180  
 GACCTTAAAC AATTGGCCCC AAGCCTCTCC TCGGAGAAAA TTGAATCAT TCGTACACT 240  
 AAGAATTGGG TTCAAAACAT TCTAAACCCA GATTTCAGCAG ATGTGAAGGA ACTGATTAAA 300  
 AAGTGGGAGA AACAGGTCAQ CCAAAAGAAA AAGCAAAAAGA ATGGGAAAAA ACATCAAAA 360  
 AAGAAAGTTC TGAAGTTGG AAAATCTCAA CTTTCTCTCT AAAGAAGAC TACAATAAG 420  
 AGCACTTCA CAAATAAGT TCTGTGTAAA AAATGTCTTA TTTAATTAT ACGCTATCA 480  
 TTCAAAGGA GGAATGGCATA TAATACAAAG GCTTATTAAT TGAAGTAAA AATTTAAGAA 540  
 ATTACTCTGA AATTGTAACT AAAGTTAGAA ATTGTATT TTGAATCCAA ACGTTAAGAA 600  
 TTGTTAAAGG CTATGATGTT CTTTGTCTT CTACCAACCA CAGTGTAAAT TCAATATGC 660  
 TTAAGGCCAT GATTTCAGTA ATACCAATGT CTACACAGAT GTTCAACAAA CCNATATCCA 720  
 CTCAACAG CTCTCTGGAA GAGCAGCCCT AGGCTTCCAC GTACTGACG CTCACAGAG 780  
 TATCTGAGGC ACATGTGAGC AAGTCTAAG CTTGTGTAGCA TCTGTGTGAG CCAAGCAAT 840  
 TGAATTTGAG CTGAGACCTCA CCAAGCTGCT GTGGCCATCA ACCTCTGTAT TTGAATCAG 900  
 CTACAGGCT CACAGCAAT GTGTCTGAGA GATTCATGCT GATTGTATT GGGTATCACC 960  
 ACTGGAGATC ACAAGTGTGT GCTTTCAGA GCTCTCTTC TGGCTTTGA AGCATGTGA 1020  
 TTCCATCTTG CCGCTCAGG CTGAACACTT TATTTCTTT TGTTCGCTT TGTCTCATTC 1080  
 AAGTCAGCTC TTCTCATCT TACCACAATG CAGTGCCTTT CTCTCTCA GTGCACTGT 1140  
 CATATGCTCT GATTATCTGT AGTCAACTCC TTTCATCT TGTGCCAAC ACCCCACAGA 1200  
 AGTCTTCTCT TCTCCCAAT CATCTCACT CAGTCCAGCT TAGTICAAGT CCGTCTCTT 1260  
 AAATAAAGCT TTTGGACAC ACAATATATC TTAAGCTCC TGTTTCACTT GGTTCAGTAC 1320  
 CAGATGGTG AACACATA CTGACACTA TCTGTGGTG TTATCTCAT CTCTCAAC 1380  
 AGATGTGAC CTCTTGAGG GCAAGAGCCA CAGTATATT CCGTGTCTT TCCAGCTGC 1440  
 CTAATAATAC TGTGGAAC TA GTTTTAATA ATTTTAAAT TGATGTGTT ATGGCGAGGA 1500  
 TGGCAACGAG ACCATTGTCT CAGAGCAGGT GCTGGCTCTT TCTGGCTAC TCCATGTGG 1560  
 CTAGCTCTGT GTAACCTCTT ACTTATTATC TTGAGGACAC TCACTACAGU GACACAGGAT 1620  
 GATCCAACAT CTTGTCTCTT TTATGACAGG ATGTGTCTC ACCTCTTCCA ACAATAGAA 1680  
 GCAGTGTGTA AAACACTCTC GGATTTCTGT GACTGTTTT AAAAATAATA CAGTTAGCG 1740  
 AAATCATAT AATCTTACAA TGAAGAGGAC TTTATAGATC AGCAGTGCAC CAACCTTTT 1800  
 CAACCATAC AAAAATTOCT TTCCCGAAG GAAAGGGCT TTCTCAATA GCTCACTGT 1860  
 TCTAAGATCT AACAAGATAG CCACCGAGAT CTTATCGAA ACTCAATTTA GGCAAAATGT 1920  
 AGTTTATTGT TCGGTTACT TGTTCAGAG TTGTGATGT GATTATCAT TACCACCA 1980  
 TCTCCATGA AGAATGAAA CGGTGAAGTA CTAAGGCTCA GAGGAAGGAC CAAAGTGGT 2040  
 TACTGGAAGC ATGATTTGTT CCAAGTATGC CTCTGACGGA TGTGGAAC TCTTCCAGG 2100  
 GGAGGTTGAG TGAATTTGT AGGAGAGGTT GTCTGTGCC AGAATTAATA CCTATACTCA 2160  
 CTTTCCCAA TGAATCACT GCTCACTGT CTGATGATT AGAGTGTCT GTCCGTGGAGA 2220  
 TCCCAACGGA AGGCTCTATC TAATCATGAA ACTCCCTAGT TCTTCACTG AACCTCCGTG 2280  
 AAAAATCTAA GTGTCTATA AATTGAGAG TGTGTGACC ACTTACTCT CATCTCAG 2340  
 GTAGACAGTA TATACTAAC AACCAAGAC TACATATTGT CACTGACACA CAGCTTATA 2400  
 TCTATTATCA TATATATACA TACATGCATA CACTCTGAAA GCAAAATA TTTCACTTCA 2460  
 AAACAGTATT GACTGTGATA CTTGTGAATT TGAATAATT TCTTGTGTA AATAGAATGG 2520  
 TATCAATAAA TAGACATTA ATCAG

Seq ID NO: 114 Protein sequence

Protein Accession #: NP\_032407

1 11 21 31 41 51  
 MKKSGVPLFI GILLVLVIG QGTPVVRKGR CSCISNTQGT IHLQSLKDLK QFAPSPCKEK 60  
 IEHATLNG VQTCNLPSDA DYKELKIKW KQVSRKQKQ NGKKHQKKV LKVRKSRSR 120  
 QKKTT

Seq ID NO: 115 DNA sequence

Nucleic Acid Accession #: NM\_003238.1

Coding sequence: 182-1426 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CAAGCAGGAT ACGTTTTCT GTTGGCAAT GACTAGATT TTGCAAAAAG TTTCGCATCA 60  
 AAAACAACAA ACAACAACAA AAAACAACAA AACCTCTCTT GATCTATCT TTGAAATGG 120

TTGATTCTTT TTGTTTATTT CTGAGTTTAA AAAACAAGTT TTTTTCAC TTTTAAAA 189  
 AATGCACTAC TTGTTGCTAG GGGTTCTTAT CTGTCACAGG TCGCGCTCAG 249  
 CTCTCTACG TCGACGACAC TGGATATGGA CCGATTCATG CGCAAGAGAG TGGAGCGGAT 300  
 CCGCGGGGAC ATCTCTGAGCA AGCTGAAAGCT CACCAAGTCCC CCAAGAGACT ATCTCGAGGC 360  
 CAGGAAAGTC CCGCCGGAGG TGATTTCAT CTACAACAGC ACCAGGAGCT TGTCTCAAGA 420  
 GAAGGGCGAG CGGAGGGCGG CCGCTTCGGA GCGCGAGAGG AGCGAGGAGA AGTACTACGC 480  
 CAAGGAGGTT TACAAATAG ACA TCGCGCC CTCTCTCCC TCGGAAAATT CCAATCGCGC 540  
 CACTTTCTAC AGACCTACT TCGAAGATTG TCGATTTTAC GTCTCCAGCA TGGAGAAAGA 600  
 TGTCTCCAAT TTGGTAAAG CAGAGTTACG AGTCTTTCGT TTGCAGAAC CAAAGGCCAG 660  
 ATGCTCTGAA CAAAGCGATT AGCTATATCA GATTCTCAAG TCCAAGATT TAACATCTCT 720  
 AACCCAGCGC TACACTGACA GCAAAGTTGT GAAAACAAGA CCGAAGAGGC AATGCGCTCT 780  
 CTGTGATGTA ACTGATCTGT TCGATGAATG GCTTCACCAT AAGACAGGA ACTTGGATT 840  
 TAAATAAGCT TCGATCTGCT CCGCTCGCAT TTTTGACCA TCTAATATT ACATGATCC 900  
 AATAAAGT GAAGAACTAG AAGCAAGATT TGCAGGTATT GATGCGACCT CCACATATAC 960  
 CAGTGTGTAT CAGAAAACTA TAAAGTCAC TAGGAAAAAA AACAATGGGA AGACCCACCA 1020  
 TCTCTGCTA ATGTTATTGC CTTCTACAG ACTTGAGTCA CAACAGACCA ACCGGCGGAA 1080  
 GAAGCGTGCT TTGGATCGG CCTATTGCTT TAGAAATGTG CAGHATAATT GCTGCTTACG 1140  
 TCACTTTAC ATTATTTCA AGAGGATCT AGGTGAGAAA TGGATACAG AACCCAAAG 1200  
 GTACAATGC AACTCTGTG CTGGAGCATG CCGTATTTTA TGGAGTTACG ACACTACAGA 1260  
 CAGCAGGCTC CTGAGCTTAT ATAATACAT AAATCCAGAA GCATCTGCTT CTCTCTGCT 1320  
 CGTGTCCCAA GATTTAGAAC CTCTAACCAT TCTCTACTAC ATTGGCAAAA CACCCAAAGT 1380  
 TGAAACAGCT TCTAATAGA TTGTAAGTC TTGCAATGC AGCTAAAATT CTGGAAAAG 1440  
 TGCCAGAGCT AAAATGACA TGAATGATG AATGATGATG ACGACACCA GATGATGCT 1500  
 TTGTAACAAG AACTATAAG GAGCTCTGCT TCAAGATCT TAAATAATT TTGAAAAGCG 1560  
 GGTACTAGT CAGACATTT GGAAGTTGT GTTCTGTTTG TAAAACTGG CATTGTACAC 1620  
 AAAAAAAGT GAAGGCGCTTA TCTCATTTT CACCTACTTT GTAAAGTGGA GAGACAAGAA 1680  
 GCAAATTTT TAAA

Seq ID NO: 116 Protein sequence  
 Protein Accession #: NP\_00329.1

1 11 21 31 41 51  
 | | | | |  
 MRYCYLSAFL ILHLTVYALS LSTCTSLMD QPMKRIRBAI RQQLSKKL TSPDEYPEP 60  
 35 EYVPEVVIS YNSTRDLQE KASRAAAACE RERSDEEYVA KEVYKIDMP PFPSENAFP 120  
 TFRYPFRIV RFDVSMABKN ASNLVKAEPF VRLQNPKR VPEQRIELY ILKSKDLTSP 180  
 TQRYIDSKVV KTRABGEWLS PDYTDVHVEW LHKHDKRNLG KISLHCPCCT PVPNSNYIP 240  
 NKSELEEARF AGIDGTSTYT SGDQKTFST RKKNNGKTPH LLMLLPSYR LESQTNRRK 300  
 40 KRALDAAYCF RNVQDNCCI R PLYDFDKRL GWKWIHEPKG YNANFCAGAC PYTWSDDTGH 360  
 SRVLSLYNTI NPEASAPFC VSRQLEPLTI LYVIGKTPKI EQLSNMIVKS CKGS

Seq ID NO: 117 DNA sequence  
 Nucleotide Accession #: NM\_000995.1  
 Coding sequence: 26-2299 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 | | | | |  
 CAGCACCCAG CTCGCCGCCA CCGCATTGGT CCGACACACC GCCTCGCTCG TCTGTCTCAC 60  
 50 CTCTGGCTGC CTCGGGCGCT CCGACAGAGG CAGAGAGGCC GTTGGGTGAG ACTGGTGGCC 120  
 CGAGATGCTT CCGGAACTGC AGGAAACCCA GCGGCGGCTG CAGAGAGTTC GAGACTGCTT 180  
 GCGGACGAG GTCAGGAGGA TCAGCTCTCT GAAACAACAG CATTGTGGAT GTGAGCGGTG 240  
 CGGATGTCCG CAGCTCAGTAC GCAACCGGCT ACCAGCGGTG CGCGCCCTGC TCCACTGTGG 300  
 GCCCGGCTTC TGCTTCCCG CGGTGGCGCT CATCCAGACG GAGAAGCGGG CGCCGTGTGG 360  
 55 CCGCTGCCCG GCGGGCTTCA CCGGCAACGG CTCGACAGC ACCGACGTCA AGAGTGTGCA 420  
 CCGCCACCCC TGCTTCCCGG GAGTCCGGTG TATCAACACC AGCCCGGGGT TCCGCTCGGA 480  
 GCTTGTCCCG CCGGATGCTG GAGTCCCGAC CCGACAGGCG GTTGGGCTGG CTTTCCGCAA 540  
 GCGCAACAG CAGGTTTTCG CCGACATCAA CCGAGTGTGAG ACCGGCCAAC ATACTGCTG 600  
 CCGCACTCCG GTTGTCATCA ACAACCGGGG CTCCTTCAGG TGGCGGCGGT GCGACCGCGG 660  
 60 CTCTGTGGGC GACCAAGGCT CCGGCTGCCA CCGCGCGGCA CAGCGCTTCT GCGCCGAGCG 720  
 CTGCGCAGCG GAGTGTCCAG AGCATGACGA CTGGGCTCTA GAGCGGATG GCTTCGGGCT 780  
 TGTGCTGTGT CCGCTTGTCT GAGCGCGCAA CCGGATCTGT TGTGGTGGCG ACATGAGCT 840  
 AGACGCTCT CCGGAGGA AGCTGGCTG CCGGAGAGCG CAGTGTCCGT AGGACAAGT 900  
 CGTGACTGTG CCGAAGTACA GCGAGGAGGA TTGGAGCCG GATGGCATCG GAGACGCTCG 960  
 CGATCCGGAT CCGCAGCGGG AGCGGGTCCC CAATGAAAAG GACAACTGCC CGCTGTGGCG 1020  
 65 UAACCCGAGC CAGCGCAACA CCGACGAGGA CAAGTGGGCG GATGGGTGCG ACAAGTCCCG 1080  
 GTTCCAGAG AGACGACACC AAAAGGACAC AGACACGAGC GCGCGGGGCG ATGCTGCCGA 1140  
 CAGGACATCT GAGCGGACCG GATCCGCGAA CCAAGGCGAC AGTGTGCTGA GGGTACCGAA 1200  
 CTCAGACCC AGGACACATG ATGGCGATGG TATAGGGGAT GCTTGTGACA ACTGTCCCA 1260  
 70 GAGAGGACAC CCGGATCAGG CCGATGTGGA CCAAGACTTT GTGGGAGATG CTTGTGACGA 1320  
 CGATCAACAG CAGGATGGAG AGCGACATGA GGAATCTCGG GACAACATGTC CCAAGTGTCC 1380  
 TAACAGTGGC CAGGAGGACT CAGACACAGA TGGCAGTGA GATGCTGTGCA ACAGACAGA 1440  
 CGACATGAC AGTACATGCT ACATGTCCGA CAACTGGGCG GTGTGTGCTA ACCCGGCA 1500  
 GCGAGAGCG GACAGGGACG CGGTGGGCGA CGTGTGCCAG GACGACTTTG ATGACAGACA 1560  
 75 GGTGTGAGAC AAGATGACG TGTGTCCGGA GAAOCTGAAA GTCAAGCTCA CCGACTGTAG 1620  
 GGGCTTCCAG CAGATGCTGC TGAACCCGGA GGGTGAACGG CAGATTGACG CCAACTGGGT 1680  
 GTGCTCAAC CAGGAGAGGG AGATGTGCA GACAATGAAC AGCAACACG CGTGGTGTG 1740  
 GGGTATACAT GCGTTCAATG GCGTGACTT CAGAGGACAC TTCCATGTGA ACAGCTGAC 1800

GGATGACGAC TATCGGGGCT TCACTTGG GTACACGAC AGCTCAGCT TCACTGGT 1860  
 CATGTGCGAG CAGATGAGG AC AACTATAT GAGCGGGAAC CCCCCTCGTG CAGTGGCGGA 1920  
 KCCTTGGCATC CAACCTAAGG CTGTGAAGTC TTCCACAGCG CCGCGGGAAC CTTGGCGGAA 1980  
 CCTCTGTGGG CATACAGGAG ACACAGAGTC CCAAGTGGCG CTGCTGTGGA AGGACCCGCG 2040  
 AAACCTGGGT TGGAAAGGACA AGAAGTCTTA TGGTGTGTT CTGACGCAAC GCGCCCAAGT 2100  
 GGGCTACATC AGGGTCCGAT TCTATGAAGG CCTGAGAGTC GTGGCGGACA GCMACTGGT 2160  
 CTGAGACACA ACCATCGGGG GTGGCGGCTT GGGGTCTCT TCGTCTGCC AGGAGAAGAT 2220  
 CATCTGGGCC AAGCTCCGTT ACCGCTGCAA TGACACATC CCAAGAGGAT ATGAGACCCA 2280  
 TGAGCTGGCG CAAGCTTAGG GACCAGGGTG AGGACCCGCC GGATGACAG CACCTCAAC 2340  
 CGCGCTGGAT GGGGGCTCTG CACCCAGGCC AAGGGGTGGC CGTCTGAGG GGGAAATGAG 2400  
 AAGGGCTCAG AGAGGACAAA AATAAGTGTG TGTGCAAGG

Seq ID NO: 118 Protein sequence  
 Protein Accession #: NP\_00036.1

1 11 21 31 41 51  
 | | | | |  
 MYPDTACVLL LTLAALGASG QGQSLGSDL GPQMLRHLQE TNAALQDVDR WLRQVQREIT 60  
 FLKNTVMBCD ACCMQQSVRT GLFSVRPLLH CAPQPFPGV ACIQTESGGR CGPCPAGFTG 120  
 NGSHCTDVNE CNAHPCFPRV RCINTSPGFR CEACPFGYSG PTHGGVQLAF AKANKQVCTD 180  
 INBECTGQHN CVPNSVCINT RGSFQCGPQ POFVGDQASG CQRGAQRFCF DGPSPBCEHE 240  
 ADCVLERLDG RSCVCRVWGA GNGLCGRDT DLDGPFDEKL RCFEPQRRKD NCVTVNPSNG 300  
 EDVDRRGDGD AGCTPADGDD VPEKDNCPV VAPFGQNTD EKKVGDAGDN CQSGNDQPK 360  
 PTDQGRKIDA GDDIDGDRPI RNSAGNPRV PMSDQKSDG DQIGDADCNQ POKSNIDQAD 420  
 VDHDVFGDAC DSDQDQDGG HQDSRDNCPY VPNSAQESDD HDGQDADCD DNDNDGVPDS 480  
 RDNCRLVPNP GQEDADRDRG GDVQDQDFDA DKVVKDKIDG PENAEVLTLD FRAFTVVLDD 540  
 PEGDAQIDFN WYVNLQGRBI VQTMNSDPGL AVOYTFNPG DFEQTHFNVT YIDDDYAGFI 600  
 FGQDSSSFY VVMWQMGBQY YWQANPFRV APBGLQLAV KSTPGQBL RNALVHTGDT 660  
 SSVRLWKD PRNVGWDKSL SYVPLQLQRP QVQYRVRVFP EGPFLVADSN VLDLTDTRG 720  
 RLGVFCFSQE NIWNLKRYR CNDTPEDEYE THQLRQA

Seq ID NO: 119 DNA sequence  
 Nucleic Acid Accession #: NM\_014211  
 Coding sequence: 157-1479 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 | | | | |  
 GGGACAGGCG TGAGGATGAG GAGAACCCTG GGGACCCAGA AGACCGTGCC TTGCGCGGAA 60  
 TCTCTGCTTG TAGGCTGAA GGAATGCCCT TACAGACGCC TCAACAACCTA CTTGTGTGAT 120  
 CTTACTTGAG CCGCTGTG TGAGCAGCTT GTCAAGCTGA ACTCAAGCTT CCACTGGGCC 180  
 TCTGTGTGTC TGAGTCTCTT CACTGAGAGG ATGTGTCATC AGGGGAGTCA GTTCAACGTC 240  
 GAGGTGCGCA GAAGTGACAA GCCTTCCCTG CTTGGCTTTG AGAACCTCAC AGCAGGATAT 300  
 AACAAATTC TCAAGCCCAA TTTTGTGTGA GAACCGGTAC AGATAGCGCT CATCTTGAGC 360  
 ATTGCAAGTA TCTCTAGCAT TTCAAGAGAT AACATGAGCT ACACAGCCAC GATTAATGACT 420  
 CGACAGCGCT GAGTGACCA GCGCCTGTG TTTGAAAGCA ACAAGAGCTT CACTGTGAT 480  
 GCGCGCTCGT TGAATGCTCT CTGGTGCCA GATACCTTACA TTTGGAGTCC CAAGAAGTCC 540  
 TGTTCCTATG AAGTCAGTGT GGGAAACAGG CTCATCGGCC TCTCTCCAA TGGCAGCGTC 600  
 CTGTATGCC TCAAGATCAC GACAACCTGT GAATGTACA CAGTATGCTGC TAAATACCCC 660  
 ATGACACAC AGACATGCAA GTTGCAGCTT CCAAGCTGGG GCGTATGATG AAATGATG 720  
 GAGTACACTT GGCTCAGG GAGAGTCACT GTGCTGAGCT TGGAAACACT GCGCTCTGCT 780  
 CAGTACACA TAGAAGCTTA TTTCACCTA CTCACAGAT CCGACAGGGA CAGAGGAAT 840  
 TACTAGATG TGTGCTTCA GTTTGAGCTT CGGAGGAATG TCTGTATT TACTTTGGA 900  
 ACCTACGTC CTCTGACTTT CCTGTGTGTG TGTGCTGGG TTTCACTTTG GATCTCTCTC 960  
 GATTCAGTCC CTGCAAGAAC CTCGATTGGA GTGAGCAAGC TGTTATCAAT GACCACACTG 1020  
 ATGATCGGTT CCGGCACTTC TCTTCCAAAC ACCAAGTCTT TCATCAAGGC CATCATATGT 1080  
 TACTTGCGGA TCTGCTTAG CTGTGTTT GGGCGCTCC TAGAATATGC AGTGTCTCA 1140  
 TACAGTCTCT TACAGCAAGT GGCAGCGAAA GATAGGGGGA CAACAAAGGA AGTAGAAGAA 1200  
 TCGAGTATTA CTAATATCAT CAACAGCTCC ATCTCCAGCT TAAACGGAA GATCAGCAAT 1260  
 GCGACGATTG AAATTTCCAG CGACAAGGTT GACTACAGTG ACTTGACAAT GAAACACGAC 1320  
 CAGACATGCA AGTTGTGCTT CCGAGAAAAG ATGGCGAGGA TTTTGATTA TTTCACAATT 1380  
 CAAACACCA GTAATGCTCA TCACTACTCT GAACTACTCT TCTCTGAT TTATGTACA 1440  
 GCAATGTTT TTACTGGG ATACAGCTG TATTTTSGA TCAATGTA ATTCTTCCA 1500  
 TGGCATAGTT CTTCACAGG ACAAGATAAT GATGTAAATG GTATTTTAGG CCAAGTGTCC 1560  
 ACCACATCCC AATGTGGCTA CAAGTGAAGT AAATAATATT TGAATCTTTC TGCTCAAGA 1620  
 ATGAAGCTCC AACCATGTT CTAAGCTGTG TAGAAGTCTT AGCATATAG GATCTGTATA 1680  
 TAGAAACATC AGTCATCTC TTTTCACTT TAATCAAGGA CATTCCCATG GACCCACAGA 1740  
 TIACAAATGT ACTGAGCTCT GTTATGTG TGCTGCTGT GTTGCTCAAT ACCTATACA 1800  
 AAGATAGGGA AGGAGACCAT TGGTAAACC TCAAGTGTCA GAAGTGTGT TCAAGTAA 1860  
 TATACATGTT TTTTACTAAA TCTGCGAGT GCTTATAAAA TACATTTGTG CCTATTGAG 1920  
 GAGTACACAT TCTGATTTT TTTTCTGTT TAAATGAAA TATGGGCTTA TGTCATTTCA 1980  
 TTGGAAGTCA ATGCATTAAC TCAATACCAA GATGAGTTT TAAATAATGA ATATTATTTA 2040  
 ATACACAAAC AGAATATGCA ATAAGTGTCA TCAATGAAA TTCAAATATA 2100  
 ATGTAGAAA AAATATAGT ATCAACATC TAAACAAAT CCTCGTGTCT AAGATACAT 2160  
 GGATTCCCCA TACTGGAAG ACCTGAGGCC TTATTCTTCC CACTATGCAT ACTGTATCA 2220  
 TTTATTTATTA TACACATCT CATCTAAAC TATACTAAAG CCTTCTCC ATGCATGAT 2280  
 GGAATGGGAA GATTTTTTT TAACTTGTC TAGAAGTCTT AATATGGCT GTTGCGATCA 2340  
 AGGCTTGCAG AATTGAGCT ATTTCTAGC TGCTTTATT CACATAGTA TGGGTGACTA 2400

AAAGTACTGG GTTGACTCAG AGAUTCCTG TCATCTTGTC ATTGCTGCTA CTTCAACACT 2460  
 GAGCAACACT CTCCTCAAGTGG CAGATCCCTG GTATCATTCG AAGAGGAGCA TGCATCCCT 2520  
 TGCTCTAATG ATCAGGAATG ATGCTTATTA GAAACAACAC TGCITGACCC AGGACCAAGT 2580  
 GGCTTAGGCTT AAGTAAACTT GGCCTTGCTC AGATCCCTGA TCCTTCACGC TGGCTGCTGC 2640  
 TATGGTGCTT ATCCCGCATG AGCAGGAGCG TGCTGGCCCT GAGTACTGAA CTTTCTGAGT 2700  
 AACAAATGAG CACGCTTACG AACCTATGTT CAGGTTGCGG GTGAGCTGCC CTCTCAAAAT 2760  
 CCAGCCAGAG ATGCAGATG CTCGCCAGAT CTCAGCCAC AGTACCAAAA GTGATTTTTC 2820  
 GTGTGCCAG GGTAAAGGCT TCAGGTTCCG CCTCATGAT TTATAGCAAT CTGCCATCT 2880  
 TTAATTCTCT AGCTTCTGCT TCTAATAAAT GCACGCGTTT ACCTTCTCT TCAGAAATAA 2940  
 AACAGGCTC TAAAGGATGA TTCCCTCTCT ATAACTCCCT AGAGCCACAG GTTCTCAATC 3000  
 CTITTCOCAT TATACTTCTC ACAATTCACT TCTATGAGT TTGATCAOCT GATTITTTTA 3060  
 TATAGGCCA TCTAATTGTT GATTITTTAA AAAAGGTTTA AAGAGATAA AATGTGGTGG 3120  
 AACCCCACTT AAGCATTTT TTTATATAAA AACAAATGATA AAGATGTGAA CTGTGAATA 3240  
 AATATACCAT ATTAGCTACC CACC

Seq ID NO: 120 Protein sequence  
 Protein Accession #: NP\_055626.1

1 11 21 31 41 51  
 MNYSLHLAFV CLSLPTERMIC IQGSGFNVBI GRSDKLSLPG FENLTAGYNN FLRPNRGPEP 60  
 VQALTLDA SSSSSSSNR DYATITLQ RWMDFQVFE GNKSPDAR LVEFLVWPT 120  
 YVYBKSXSL HEVTVNRLH KLPNGVLTLY ALRITTVAC NMDLSKPYMD TQTKLQLS 180  
 WGYDNDVFE TWLRGNDVR GLBHLRLAQ TIERFYPLTV RSQBTGNYT RLVLQGLCT 240  
 NVLYFILBTY VPSTFLVLVS WVSFWISLDS VPARTCIQVT TVLSMTLMI GRTSLPNTN 300  
 CFKADIVLY QICPSFVPGA LLEYAVHYS SLQMAAKDR GTTKEVEYS ITNINSSIS 360  
 SFKRKISFAS IEISSDNVDY SLDMTKTSK FKFVFRKMG RVDYFTQN PSNVDEHYSKL 420  
 LFLPLMLAN VFYWAYFYMYT

Seq ID NO: 121 DNA sequence  
 Nucleic Acid Accession #: NM\_001854  
 Coding sequence: 163-5582 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 AACCATCAAA TTTAGAAGAA AAAGCCCTTT GACTTTTTTC CCTCTCCCT CCCCAATGGC 60  
 TGTGTAGCAA ACACTCCCTG GCATCACTTG GAAAGGAGCA AGTTGTGCTG CAGTCGCAAT 120  
 TTCTGGTTT GAGTTCACAG TGTCTAGTGC GAGCTGCTGA GATGGAGCCCT TGGCTCTTA 180  
 GGTGGAAAC GAAACCGCTG CTCGKGATT TCACGCTAAC AACCTCGCA TGAACCTTC 240  
 TCTTCCAAGC TAGAGAGGTC AGAGGAGCTC CTCACGTTGA TGACTAAAA GCACATAGAT 300  
 TTCACAATTC TCCAGAGGGA ATATCAAAAA CAAAGGGGATT TTGCACAAA ACAGAAAGATT 360  
 CTAAAGGCTC AGATACTGCT TACAGAGTTT CAAAGCAAGC ACAACTCAGT GCCCCACAAA 420  
 AACAGTTATT TCCAGGTGGA ACTTCCGAG AGACTTTTC AATACATATT ACAGTAAAC 480  
 CAATAAAGG AATCGAGCTT TCCCTTTTAT CTATATATA TGAGCATGCT ATTCAGCA 540  
 TTGTTGTGA GGTGGGAGA TCACCTGTTT TCTGTTTGA AGCCACACT GGAAAACCTG 600  
 CCCCAGAGA CATCTCCCTC TTCAGAACTG TTAACATCGC TGACGGGAG TGCCATCGGG 660  
 TAGCAAGCA CGTGGGAGA AGAAGCTGTA CAATGATTTG TGAATGTAA AGAAAACCA 720  
 CGAAACCACT TGATAAGAG ACAATAGCAA TTGTTGATAC CAATGGAATC ACGTTTGTG 780  
 GAAACAAGAT TTTATGAGA GAAGTTTGT AGGGAGACAT TCACGACTTT TGTATCAAC 840  
 GTGATCCAA GCGACGATAT GACTACTGTG AGCAATTATG TCCAGACTGT GACTCTTCAG 900  
 CACCCAAGGC TGCTCAAGCT CAGGAACCTC AGATAGATGA GTATGCACA GAGGATATA 960  
 TCGAATATGA CTATGAGTAT GGGGAAGCAG AGTATAAAGA GCTCTAAAGT GTACAGAGG 1020  
 GACCCACTGT AACTGAGGAG ACAATAGCAC AGACGAGAGC AAACACTGTT GATGATTTTC 1080  
 AAGAAATCAA CTATGGAACA ATGGAAGATT ACCAGACAGA AGCTCCFAGG CATGTTTCT 1140  
 GGCACATGA GCAAAATGCA GTTGAAGAAA TATTACTTGA AGAATATCTA ACGGGAGAGG 1200  
 ATATGATTTC CCGAGAGAAA AATCTGAGG ATACACTATA TGAATAACAA GAATAGACG 1260  
 GCAGGGATTG TGATCTCTG GTAGATGGAG ATTAGGCCA ATATGATTTT TATGAATATA 1320  
 AAGAAATGA AGATAAACCA ACAAGCCCTC CTAATGAAGA ATTGTGCA GBTGTACCG 1380  
 CAGAAACTGA TATTACAGAA ACAAGCATAA ATGGCCATCG TGCATATGGA GAGAAAGGAC 1440  
 AGAAAGGAGA ACCAGAGTGT GTTGGAGCTG GTACTCTGT GGAAGAGCA CAGAGCATC 1500  
 CAGGACTGCG AGGTATTATG GGTCTCCAG GTCTACAAG CCCCAGTGA CCCCCTGGTG 1560  
 ACCCTGGGCA TAGGGGCCCC CAGGAGCTG CTGGCTTACC AGGGGCTGAT GGTCTACCTG 1620  
 TGCTCTCTG TACTATTTTG ATTATTACGT TCGGTATG TGTTGATGGT TCCAAAGGAC 1680  
 CAACCATCTC TGCTCAGGAA GCTCAGGCTC AAGCTATCTT TCAGCAGAGCT CGAGTTGCTC 1740  
 TGTAGAGCCC ACTGTGGCCA ATGCTGTCTA CTGGAAGAC AGCTCTCTGT GGGGGGCTG 1800  
 GTTCACTCTG GGCACAAAGT GAGATGTGTG ATCCAGGCTC TCAGGGCCCT CAGAGCGTCC 1860  
 AGGCTCCCC TGCTCAACG GGAACAACCT GAAAAAGGGG TGCTCCAGGT GCAGATGGAG 1920  
 GAAGAGGAAT GCGAGGAGAA CTTGGGGCAA AGGAGATCG AGGTTTGTGAT GGACTTCGG 1980  
 GTCTGCCAGG TGACAAGGTT CACAGGGGTG AACGAGGCTC TCAAGTGCTC CTAGTGCTCT 2040  
 CTGTTGATGA TGGAA TGAGG GGAGAGATGT GAGAAATGTG ACCAAGAGTT CTTCAGGCT 2100  
 AAGCTGGCC ACTGAGTTGT CTGGGTCTCA GGGGAATCG AGGAGCTTCA CGGCGAGCT 2160  
 GTATTGCCAG TGTAGATGCG CCCCAGGAC CAAAAGGGAA CATGGGTCCC CAAGGGGAGC 2220  
 CTGGGCTCC AGGTCAACAA GGGAAATCAG GACTCTCAGG TCTTCTGTT CCACAAGCT 2280  
 CAATTGTTCC TCTGTGTGAA AAGGACCCAC AAGGAAACCC AGGACTTGCT GACTCTCTC 2340  
 GTGCTGATGG GCTCTGCTGT CACTCTGGGA AAGAAAGGCC GTCTGGAGA AAGGGGCTC 2400  
 TGGTCCCCC TGCTCAGAA GGTCTATTG GATNNCGGG CCCCAGGGA GTAAAGGAG 2460

CAGATGGTGT CAGAGGCTCT AAGGAGTCTA AAGGTGAAA GGTGGAAGAT GGTITTCAG 2520  
 GATTCAAGG TGATATGGT CTAAAGGTTG ACAGAGGAGA AGTGTGTCAA ATTGGCCCAA 2530  
 GAGGQNAAGA TGGCCCTGAA GGAACCAAAG GTGAGCAGAG CCAACTGTGA GACCCAGTCA 2540  
 CTTCAAGGTCA AGCAGGAGAA AAGGGAAAC TTGGAGTTTC AGGATTAOCA GGATATCCAG 2700  
 GAGACAAAGG TCCAAAGGGT TCCACTGGAT TCCGTGGTTT TCCAGGTGCC AATGGAGAGA 2760  
 AAGGTGCACG GGAAGTAGCT GCAAAACCAAG GCCCTCGGGG TCAGCGTGGT CCAACGGGTC 2820  
 CTCAGGTTTC AAGAGGTCCA AGAGGTCCCA CTGGGAAAC TGGGCGCAA AGGACCTTCA 2880  
 GTGGGATGG CCTCTCTGCG CTTCAAGGTT AAGAAGGTTCT TCAAGGACCT CAGGTTCCAG 2940  
 TTGGATTCCC TGGACCAAAA GGCCCTCTGT GACCACCAAG AAGGATGGCG TGCCACGAG 3000  
 ACOCTGGCCA ACOTGGGGAG ACTGGATTTC AAGGCAAGAC CGGCGCTCTT GGCCACAGGG 3060  
 GAGTGGTTGG ACCACAGGGA CCAACCGGTG AGACTGGTTC AATAGGGGAA CGTGGGTATC 3120  
 TCTGGTCTCC TGGCCCTCTT GTGTAGCAAG GTCTTCCTGG TCTCGACAGA AAGAAGGTTG 3180  
 CAAAGGGTGA TCCAGGTGTT CAGGTATCT CAGGGAAGA TGGACACGA GGAATACGTG 3240  
 GTTTCACAG GGAAGAAGGT CTCTCTGAGC CTCAGGGTGC ACCTGGACTG AAGAAGAGGG 3300  
 AAGTCCCCA GGGCCACCA GTTCCAGTTG GCTCACAG AGAAGCTGGG TCAGCGAGTA 3360  
 CAGCTGGCCC AATTGGTTTA CGAGGGGGCC CGGACCTTCA GGGTCTCTCT GTTCAGGTTG 3420  
 GAGAGAAAGG TGCTCTCGA GAAAGAGGTC CCAAGGGGCT TCAGGTGAGA GATGAGTTTC 3480  
 AAGTCTCTGT AGCTCTCTT GTGACAGTGT GTCTCCCGG CTCCCTGGGG GAAGAAGGAG 3540  
 ACAAGGGTGA AATTGGTGA GCGGGACAAA AAGGCAGCAA GGTGGCAAG GGAAGAAATG 3600  
 GCGCTCCGG TCCCCAGGT CTTCAGGAC CAGTTGGTGC CCTGGAAAT GCTGGAGGTTG 3660  
 ATGGTGAACC AGGTCTTGA GACAGCAGCG GGATGTTTGG GCAAAAAGGT GATGAGGGTG 3720  
 CCAAGAGGCT CCTGTGACCT CTTGTCTCAA TAGTCTTCA GGGTCTGCA GGCCACAGCT 3780  
 GTGAAAAGAG TGAAAATGG GATGTGTGAT CATGGGGGCG ACCTGGTCTT CAGGCCCAA 3840  
 GAGGCCCTCA AGTCTCTCA GTGAGTGTAT GACCAAGAG ACCCCAGGT TCTGTGGTT 3900  
 CAGTGTGTGG TTTTGGAAA AAGGTTGAAC CTGGGAAGC AGGAACCCA GGCCCTCTCT 3960  
 GGGAGACAGG GTTAGGGGGT CCAAGAGGAG AAGAAGGAGA GAAAGGGGAA GCTGTTCCAC 4020  
 CTGAGGCTCG TGACCTCCA GTGCCAAG GGC0CCAG TGATGATGCG CCTAAGGGA 4080  
 ACCCGGTTCC TTTTGTGTT CTTGGAGATC CTGGTCTCC TGGGGAAGT GGCCCTCGAG 4140  
 GTCAAGATGT TGTGTGTGT GATGGGTG AAGATGAGA TCTGTGTCAA GGTGCTCTCT 4200  
 CTGGCCATC TGTGTAGGCT GGCCACCAG GTCTCTGGT AALACGAGT CTTCTGGTAG 4260  
 CAGGAGGTGC AGAGGGAAGA CAAAGTGAAA AAGGTCTAA GGGGGAAGCA GTTGCAAG 4320  
 TCTCTCTGG AAAAACCGCG CCAGTGGGTC CTCAGGAC: TGACGGAAG CTTGGTCCAG 4380  
 AAGGCTCTGG GGGCTATCCT GTGCTCTGG GAGACAGAG GTCTCTGGA GCTCGAGGCT 4440  
 AAGATGAGAC AGGCTCTCTT GGTGACCTGT CAGCTCTACC TGTGTCAA GGTGCTCTCT 4500  
 GCTTCAGGTT TGAAAAGGA CATCTCTGTT TAATGGGCT GATTGGTCT CAGGAGAAC 4560  
 AAGGGGAAA AGGTGACCGA GGGCTCCTG GAACTCAAG ATCTCAGGA GCAAAAGGG 4620  
 ATGGGGAAAT TCTGTGCTCT GCTGGTCTT TAGGTCCACC TGGTCTCCA GGTCTACAG 4680  
 GTCTCAAGC CCAAAAGGCT ACAAAGGCT CTACTGAGAC CGTGGGCCA AAGAGTGACA 4740  
 GTGTCTCTC AGGCGCTCTT GGGCTCTGAT GACACAGTGT TAGAGTCAAT CAGCTTAC 4800  
 CAATTTTGT CTCGCAAAA ACAGAGAAC ATACTGAAG CATGCAAG TACGATAGT 4860  
 ATAATATCT TGATTACTGT GATGGAATGG AAGAATATT TGTTCCTCT AATTCCTGTA 4920  
 ACAAAGACAT CGAGCATATG AAATTTCCAA TGGGTACTCA GACCAATCCA GCGGCACTT 4980  
 GTAAAGACCT GCAACTCAAC CATCTGACT TCCAGATGG TGAATATTGG ATTGATCTTA 5040  
 ACCAAGGTG TCTCAGGAGT TCTCTCAAG TTACTGTAA TTTCACAT GGTGGTGAGA 5100  
 CTGCAATTA TCCAGACAAA AATCTGAGG GAGTAAGAT TTCACTATG CCAAGAGGA 5160  
 AACCGAAGG TTGTTTAAAT GAATTTAAGA GGGGAAACCT GCTTTCATC TTAGATGTTG 5220  
 AAGGAAATTC CATCAATATG GTGCAAAATG CATCTCGAA ACTTCTGACT GCCTCTGCTC 5280  
 GGCAGAAATTT CAGCTACCAC TGTCTACAT CAGCAGGCTG GTATGATGTG TCATCAGGAA 5340  
 GTTATGACAA AGCACTTCGT TCTCTGGAT CAAATGATGA GGAAGTGTC TATGACAATA 5400  
 ATCTCTTTAT CAAACACGTG TATGATGGT GTACGTCAG AAGAGCTAT GAAGAACTG 5460  
 TCAATGATAT CATACACCA AAATTTGATA AAGTACCTA TGTGATGTC ATGATCAGTG 5520  
 ACTTTGGTGA TCAGAAATCA AAGTTGGAT TGAAGTTGG TCTGTTTGT TTTCTTGCTG 5580  
 AAGATTAAAG CAAAGAACAT ATCAAAATCA CAGAAAATGT ACCTTGTGTC CACCAACAG 5640  
 TTTTGTGCA CATGCAAGTT TTGAATAAG ATGTATGGAA AACAAAGCTG CATATACAGG 5700  
 TACCATTIAG GAAATACCCA TCCCTTTTGT GGGGCAAGAT CACACAGCA ACTCTTGAAA 5760  
 ARCAATAAG TATCATTTGG TGTGGTATG ATGAAACAG GGCTGATCT TGAITGCCAA 5820  
 TTTCTCAACT ATGCTTTTCT ATTGAATTT CTITGTGTC GTAGAAACA AAAAAAGAA 5880  
 AATATATATT CATAAAAAAT ATGGTGCTCA TTCTCATCCA TCCAGGATGT ACTAAAGAC 5940  
 TGTGTTAAAT AAATGTAAAT TATTTTGTG ACAATTCTAT ACTGTATCT GTGTCCATT 6000  
 CCAAAACTGT CAGCTGTCCC TGAATTCCG TACTCTAAT TTATAGGAT GCGCAACTCT 6060  
 GATGCAATA ATATATGAT TATGAAAAGT AAGTATGAT TTCCGATGAC CTAAGTCCD 6120  
 TTTCTTGT TAAATGAAA ATTCTTGT GTGTGTT

Seq ID NO: 122 Protein sequence:  
 Protein Accession #: NP\_001845

1 11 21 31 41 51  
 MEPWSSRWKTKRWLWDPTVT TLALTHLPA REVRGAAPVD VLKALDFHNS PEGISKTGIF 60  
 CTRNRNKRGS DTA YRVSKQA QLSAPTKQLF PGITPIPHFS ILTVPKPKG IQSFLSTFN 120  
 EHKIQQKVS QVRSVAPDF DITGAPEDF YLPRTVNA DGKWIHWAS VRKKTVTIMV 180  
 DKCKKTKPL DRSEARIVDT NGITVETRI LDEEVEFGDI QQLFIDGPK AAYDYCHYS 240  
 PDCDSSAPKA AQAQEPQIDE YAPDEIIEYD YHYGEAEYKE AENVTEGPIV TBETIAQTEA 300  
 NIVDFDQEVN YGTMSYSQTE APRHVSQTN ENPVEEFTE EYLTEDEYDS QRKNSIEDTLY 360  
 ENKSIDGRDS DILVDGDLGE YDFYBYKVE DKTPSPNEB NGPQVAETD IETTSINGHC 420  
 AYGKQKQGE PA VYVPGMLV EGFPGAGPA GIMQFPLQG FTGPGFGRD RGFGRPLP 480

GADOLPGPPG TMLMLPFRYG GDGSKOPTIS AQAQAQAIL QQARIALRGP PGPMGLTGRP 540  
 GPVGGGSSG AKGSGPGRPG GPFGKPGKSG RFGADGKRGM PGPGAKGDR 600  
 GFDLPLKLG DKGHRERGPG QDPFPDPDD GMRGEDGEIG FRGLFGEAGP RLLGLPRGTP 660  
 GAGPQPMAG VDPGPFGKGN MGPGQBPFGP GGGGNPGFGP LFGPGFPGP PGBKPGQKGP 720  
 GLAGLPGADG PPGHGPBGQ SGEEKGALGPP GPQGGPKXGP PRGVKGADGV RGLKSGKOEK 780  
 GEDGPPGFGG DMGLKGDRGE VQVQIPRGXD GPFGPKGRAG PTUDPFGSQ AGHKGKLVF 840  
 GLPGYFGRQG PKGSTGFPFG FQANCKEGRAR GVAKGFPFGG QRGTPGRRG RGAAGTKCP 900  
 GPKGTSGGDG PGGVGGKPG QDPQGPVGP GPKGPFPFG RMGCPHFGQ RKGTFPQK 960  
 QPFGPGVVG PQPFTGETGP IGERGYPGFP GPFGGQLFG AAKGEGAKGD PQFGSGDK 1020  
 GPAGLRFPFG ERGLPGAQGA PGLKGGBGP GPFGVGSFG ERGSAGTAGP IJLGRGPPO 1080  
 GPPGPAEGKG APBEGKPGQG AGRDGVQGVV GLPQAPGAPG SPGEDGDKGE IGEPPGQKSK 1140  
 GKGQENPPFG PFLGQPVGA PGIAGGDGEP GPRGQKMGPG QKGDGKARGP PFPFGIQL 1200  
 GLPGFPFKG ENKDPGPGP PGPFGPRPG QPMAGDPQG PFGSVGSVG VGBKGEGEA 1260  
 GNPFGPGGAG VGGKPGERGE KGEAGPPGAA GPPKAGKPGP DDGPKGNPGP VYFPGDPGP 1320  
 GBLGPAQDQG VGGDKGDDGP FGQPPPPPS GEAGFPFGP KRGPFGAAGA EBRQGEKGA 1380  
 GEAGABGPPG TPGTVPGQGP AGKPGPBGRL GPGPVGEGG LPGAGQDGP PGPMGPPGLP 1440  
 GLKGDGSKG EKHGPGILGL IGPPGEQBGK GDRGLPTQG SPKAGDGGI PGFAPLOPP 1500  
 GPPGLPGPG PKUNKGSTGP AGQKDGSLP GPFGPPPGP EVQPLPLS SKKTRHTEG 1560  
 MQADADNLL DYSDGMEIFG GELNSLAKQI EHMKPMGTQ TNPARTCKL QLSEHPDVG 1620  
 EYWDPNQGC GSDKFSYVCN FTSGGETCIY PDKSGBVRI SSWPEKPGS WFSFKRGL 1680  
 LSYLDVBONS DMVQMTEFL LITASARQNF TYBHQSAAW YDVSSGSYDK ALRFLGNSDE 1740  
 RSMYDNNPFI KTLYDGCTSR KGYEKTBIH NTKIDQVPI VDMMSDFGD QNKQGFVEG 1800  
 PFCFLG

Seq ID No: 123 DNA sequence  
 Nucleic Acid Accession #: NM\_015886  
 Coding sequence: 485-1261 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 GAATTC<sup>1</sup>CCCC CCCCC<sup>11</sup>CCCC TCAC<sup>21</sup>TTGGT TGTCTATATG TCTGCGAGAC ATTATCAGCA 60  
 CATCTCTGT TGTTACCTGT GATTCATTT TTTCTCCTC TCCAGGTGAA TTTCATTCG 120  
 TGA<sup>31</sup>AAATTC CCAC<sup>41</sup>TAAAA TATGCAAGTAA TATATTTGT GGTTCAGACA TTTGGGCAAA 180  
 ATGTTTCACA TCTATT<sup>51</sup>TGAG GGTTAGTGCT CATCTGTGT ATTTTCTCT GCTATACAA 240  
 GTTCTCTTA GGGGTCTGCC TCATGACACT AAAAAGTAA TAGAGATTCT ACCTAGAGTT 300  
 ATCTCTAGG CTTAGATTCA ACATTTGTTT GGATTTTGA AGAAAGTCAA ATCAAGAAT 360  
 GCTCCAAAT GATGTCTTTG TAAATTCATA CCTCTGGCC CTA<sup>401</sup>TTTTT TCATGAGCC 420  
 GACTCTACC TTTCTGCTT AAGAAGAAGT AACTCTGGT GCTCTCTCT CTTCCACCCCT 480  
 CAA<sup>491</sup>AAATATA GCAATCTCT CCGTACGAG TGCATCTCT TCTCTCTCT TCTGTGAAGC 540  
 AAGTACGCT TACTCTCA CTATTCAGCA GCACTCCGC CCACCAATA ATTCTACTGA 600  
 TATTGAAGCA GCTCTGAAG CACAATTAGA TTACGGGAT ATCCCCAAG CCA<sup>601</sup>GGCGAA 660  
 GCGTACATT TCGCAAGAT ACATGATGCG CATCTCTGAT TATCATAATC AAGTTCGGGG 720  
 CAAAGTGTTC CCACCGGCAG CAAATATGGA ATATATGGT TGGGATGAAA ATCTTGCAAA 780  
 ATTCGAGAG GCTTGGGGG CTA<sup>801</sup>CTTGAT TTGGACCAT GGACCTTCT ACTTACTGAG 840  
 ATTTTGGGC CAAATCTAT CTGTACAC TGAAGATAT GCGTCTATC TCGATGTG 900  
 CAACGATGG TATGATGA GCAAGATTA TGTCTTCCA TATCCCGAGG ATTCGACCC 960  
 CAGATGTCT ATGAGATGTT TGTGTCCAT GTGCACACAT TATACGCAGA TGGTTTGGC 1020  
 CACTTCCAAT CGGATAGGAT GCGCAATTC TGTCTGCCA AACATGAATG TTTGGGATC 1080  
 TGTGTGGGA CGTGCAAGTT ACTTGTGAT CCAACTGGG CCA<sup>1101</sup>AGGGCA ATTGAGATTG 1140  
 AAGAAGCACA TATAAAGTAG GGGTACCAT TCTACTTGT CCTCAAGT ATGSGGAGCT 1200  
 TGTACTGAC AAGTGTGTT TCCAGAGG TACGTCAAC TACCTGTACT GTTATATA 1260  
 GTTATCTAT TCTCTCCAG AATATAATG ATTCTGGGA ACATGGCAT GATATATAT 1320  
 ATATGGAGAG AGAATTGTC ACATATATA CATATTTGT GCTAATCTTG TTTTCTCT 1380  
 AGTATCTCT TGTATAAAT AGTGTGTTG TAGCATGTT GTTATACTC TTGGGAATTC

Seq ID No: 124 Protein sequence  
 Protein Accession #: NP\_055970.1

1 11 21 31 41 51  
 MIAISVSSA LFLSLCLAS TVLLNSTDS SPTNNFTD EALKQLDS ADIPKARRR 60  
 YRSQNDMAI LDYHNVGRK VEPANAMEY MWVDENLKS ABAWAATCW DRGFSYLLR 120  
 LKQNL<sup>31</sup>SVRTG RYRSLQLVK PWYDVKDYA FPYQDCNFR CPMRCFPMC THYQVMYWA 180  
 SNRGICAHQ CQNMNVWGSV WRRVYLVCH YAPKGNWIGE APYKGVFCS SCTPSYGGSC 240  
 TDNLCPGVT SNLYWFK

Seq ID No: 125 DNA sequence  
 Nucleic Acid Accession #: NM\_001793  
 Coding sequence: 54-2543 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 GCGGAACACC GCGCCGCGCT CGCGGCAGCT GCTTACACCC TCTCTGCA GCCATGGGCG 60  
 TCCCTCGTGG ACCCTCTGGG TCTCTCCYCC TTCTCCAGTT TTGCTGGCTC CAOTGCGCG 120  
 CCTCGAGCC GTCCCGGGCG GTCTTCAAGG AGGCTGAAGT GACCTTGGAG CGGGAGAGCG 180  
 CGGAGCAGGA GCGCCGCCAG CGGCTGGGGA AAGTATTCAT GGGCTGCCCT GGGCAGAGC 240



CAGCTCTGTT TAGCACTGAT AATGATGACT TCACGTGCG GAATGGCGAG ACAGTCCAGG 300  
 AAAGAAGGTCT ACTGAAGGAA AGGAATCCAT TGAAGATCTT CCAATCCAAA CGTATCTTAC 360  
 GAGACACAAA GAGAGATTGG GTGGTTTGCT CAATATCTGT CCGTGAAAAA GCGAAGGGCT 420  
 CTTCCGCCCA GAGACTGAAT CAGCTCAAGT CTAATAAAGA TAGAGACACC AAGATTTTCT 480  
 ACAGCAATCA CAGGACCGAG GCAAGACAGC CCGCTGAGG TGTCTTCCCT GTAGGAAGG 540  
 AGACAGGCTG GTTGTGTTTG AATAAGCCAC TGGACCGGGA GGAGATTGGC AAGTATGAGC 600  
 TCTTTGGCCA CGCTGTGTCA GAGAATGGTG CCGTACATGGA GGACCCCATG AACATCTCOA 660  
 TCATCGTGAC GACCCAGAAT GACCACAAGC CCAAGTTTAC CCAGGACACC TTCGAGGGGA 720  
 GTGTCTTAGA GGGAGTCTTA CCAGGTACTT CTGTATGTGA GTGTACAGCC ACAGATGAGG 780  
 ATGATGCCAT CAGTACTGAT AATGGGTGGT TGTCTTACTT GATCCATAGC CAGAGACAAA 840  
 AGGACCCACA CGACCTCTAG TTCAAACTT CCGAGGACAC AGGCACATCA AGCGTCACTT 900  
 CCAATGGCCT GGACCGGGAA AAGTCCCTGT AGTACACACT GACCATCCAG GCCACAGACA 960  
 TGGATGGGGA GCGCTCCACC ACCACGGCAG TGGCAGTAGT GGAGATCCTT GATGCCAATG 1020  
 ACAATGCTCC CATGTTTGGAC CCGCAGAAGT ACGAGGGCCA TGTGCTGAG AATGCAATGG 1080  
 GCGATGAGT CAGAGAGGCT AGCGTCACTG ATCTGAGCC CCGCAACTCA CCAAGCTGGC 1140  
 GTGCCACCTA CTTATCATGT GCGGTGAGC ACGGGACCA TTTACGATC ACCACGACC 1200  
 CTGAGACAAA CCAGGGCATC CTGACAACCA GGAAGGGTGT GGATTTTGAAG GCCAAAAAAC 1260  
 AGCACACCTT CTACGTGTA A GTGACCAACG AGGCCCTT TTGTGCTGAAG CTCGCAACCT 1320  
 CCGACGCCAC CATAGTGGTC CAGGTGGAGG ATGTGAATGA GGCACTGTGT TTTGTCCAC 1380  
 CTTCCAAAGT GGTGAGGTC CAGAGAGGCA TCCCATCTGT GGAGCGTGT TGTGTCTACA 1440  
 CTGACAGAA CCTGAGAGG GAGATTCAAA GATGACATA CCGCATCTGT GAGAGCCAG 1500  
 CAGGTTGGT AGCCATGGAC CCAGACAGTG GGAAGTCTAC AGCTGTGGGC ACGCTCGAAC 1560  
 GTGAGGATGA GCAGTTTGTG AGGAACAACA TCTATGAAT CATGTCTGT GCCATGGACA 1620  
 ATGGAAGGCC TCCACCACCT GGACCGGAAA CCGTCTCTGT AACACTGATT GATGTGACAG 1680  
 ACCATGGCCC AGTCCCTGAG CCGCGTCTGA TCACCATCTG CAAOCAAAGC CCGTGGGCCG 1740  
 ACCTGTGAAA CATCACGAG AAGCGTCTGT CTTCCGACAC CTGACCTTTG CAGGCGAGC 1800  
 CTGACAGATA CTCACCTG TACTGGAGCG CAGAGATCAA CAGGAGAGT GACACAGTGT 1860  
 TCTTGTCTCT GAAGAAGTTC CTGAAGCAGG ATACATATGA CGTGCACCTT TCTCTGTCTG 1920  
 ACCATGGCAA CAAAGAGCAG CTGACGGTGA TCAGGGGCCAC TGTGTGCGAC TGCCATGGCG 1980  
 ATGTGCAAA C TGCCCTGGA CCGTGGAAG GAGGTTTCA CTCCCTGTGT CTGGGGGCTG 2040  
 TCTGGCTCT CTGTGTCTCT CTGCTGGTCT TCGTTTGT GTGTGAAGAG AAGGCGAAGA 2100  
 TCAAGGACC CCGTCAAGAT CACAAGATGT ACACCGTGA CAAGCTCTTC TACTATGGC 2160  
 AAGAGGGGGG TGGCGAAGG GACCAGGACT ATGACATCA CCGAGTCCAC CCGAGTCTTG 2220  
 AGCGCCAGGC GGAGTGGTGT CCGGCAATG ACGTGGCCAC AACCATCATC CCGACACCCA 2280  
 TGTACCGTCC TAGGCCAGCC AACCAGATG AATGGGCAA CTTTATAAT GAGAACCTGA 2340  
 AGCGGCTCAA CAGAGACCC ACAGCCCGCC CTAACACAC CTTCTTGTTG TTGAGTATG 2400  
 AAGGACGCG GCGTCCCTGA GCGTCCCTGA CTTCTCCAC CTTCTCCGCC TCGACCAA 2460  
 ACGAATATA CGATTATCTG ACGAGTGGG GCGAGCGTCT CAAGAAGCTG CAGACACTGT 2520  
 ACGGTGGGG GGAGACGAC TAGGCGGCT CCGTGCAGG CTGGGGACCA AAGCTCAAGC 2580  
 CACAGAGCAT CTCGAGGGG TCTCAGTTC CCGTTCAGT GAGGACTTC GAGCTTGTGA 2640  
 GGAAGTGGCC GTAGCAACTT GCGGAGGAGA GCGTATGAGT CTGAGTTAG AGTGGTCTCT 2700  
 TCGTTAGCTT TTAGGATGG AGGAATGGTG GCGATTTGAC TCGACACTG AAAACCTCTC 2760  
 CACCTGGGCC ACGGTTCCCT CAGAGGCCAA GTTTCGAGA GCGCTTACCTTGGCTTAAA 2820  
 TGTCTAACCC TGTGTTCTGG CCGTGGGCGT CCGTGTGACT ACCTACAGTG GACTTCTCT 2880  
 CTGGAATGGA ACCTCTTAG GCGCTCTGGT GCAACTTAAT TTTTTTTTT AATGTATCT 2940  
 TCAAAACGTT AGAGAAAGTT CTCAAAAGT GCAAGCCAGA GCTGCTGGGC CCACTGGCCG 3000  
 TCTGTCAATT CTGTGTTGCA GACCCCAATG CCGTCCATTC GGATGGATCT CTGGGTTT 3060  
 ATACTGAGTG TGCGTAGGTT GCGCTTATT TTTATTTTC CCGTGTGGT TGCTATAGAT 3120  
 GAAGGTGAG GACAATGTT TATATGTACT AGAAGTTT TATTAAAGAA A

Seq ID NO: 126 Protein sequence  
 Protein Accession #: NP\_001784

1 11 21 31 41 51  
 MGIPRGLAS LLLLVQWVLC CAASEPCRAV FREAEVILEA GGAEQEPQGA LGKVFMGCPG 60  
 QEPALFSTDN DDFTVRNGET VQERRSLKER NPKIFPSKR ILSRHKRDVY VAPISVPING 120  
 KIPFFQLNLQ LKSNKRDITK IPYSTGPGA DSPHGVFAY EKETGWLLN KPLDRBEAK 180  
 VELRGVAVSE NGASVDEPM ISIVDTGAE HKKFKTQPT KSVLGRVLP GTSVMQVTA 240  
 DEIDAITYN GVVAYSIHQ EKPDHDLMF THIRSTGTS VISSLDRK VPEYTLTQA 300  
 TMDMGDGSTT TAVAYVEILD ANDNAPMFP QKYVAHVPA AVGHEVQRLT VTDLDPNPS 360  
 AWRTATYLDG GDDGDFHTT TPESHNQGL TRKGLDFEA KNQHTLYVEV TNEAFVFLK 420  
 FISTAIVHV VBDVNEAPVF VPSKVVEVQ EQIPITGPVC VYTABDPKSE NQKYSRILR 480  
 DPAGWLAMPD DSGQVAYGT LDRBQGFVR NNHYEVNVL A MDNSPPTTG TGILLTLID 540  
 VNDHGPVPEP RQITCQNSP VRHLNLTDK DLSRHTSPQ AQLTDDSDY WTAEVNEBEG 600  
 TVVLSLKFLK KQDTYVHLS LSDHGNKBQL TVIRATVCDG HGHVETCPG WKGGFLPLVL 660  
 GAVLALLFLF LVLILLVRKK RKIKPELLLP EDDTRDNVY VYBGGGGBED QYDITQLHR 720  
 VLEGAPVVL RNDVAPTIIP TPMYRPRPAN PDEGNFIB NLKAA NTDP T APFYDITLFF 780  
 DYBGSQSDAA SLSSLTSSA DQDQYDYLN EWGSRFKLLA DMYGGGEBD

Seq ID NO: 127 DNA sequence  
 Nucleic Acid Accession #: NM\_003256.1  
 Coding sequence: 60-734 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CTCTGCTGGG CCGTCAAGTC CCGCAGAGCT CACAGGCTCA GTCCGCGATC TGCAAGTGTCA 60

TTGCTGGGAG CCTCTGGGCC GGGCCAAAGT GGTGTCTGTT GCTGCGGCTG CTGGCGTTGC 120  
 TTGGGCCCCC GGGGCTGGGT GAGGCAATGA GCTGGCCCCC GGGCCACCCT CAGCAGCACA 180  
 TCTGGCACTC GGCACATTGTG ATTGGGGGCA AAATCTCCAG TGAGAAAGTA GTTCGGGCA 240  
 GTGCAGACCC TGCTGACACT GAIAAAATATG TCCGGTATGA AATCAACAG ATAAAGATGT 300  
 TCAAAGGTTT TGAGAAAGTC AAGGATGTTC AGTATATCTA TACGCTTTT GACTCTTCC 360  
 TGTGTGTGTC GAAGCTAGA GCGAAGACCT AGAAGACATA TCTTTTGA CT GTTCAGGTCC 420  
 TCAGTGAATG AAAAGTCTTC ATCCATCTGT GCAACTACAT CGAGCCCTGG GAGGACCTGT 480  
 CCTTGGTGCA GAGGGAAGT CTGAATCATC ACTACCATCT GAACCTGGGC TGCCAAATCA 540  
 CCACCTGCTA CACAGTACCC TGTACCATCT CGGCGCCTAA CGAGTGGCTC TGGACAGACT 600  
 GGCTGTTTGA ACAGAAAGCTC TATGGTACCC AGGCTCAACA TTAATGTCTG ATGAAGCATG 660  
 TTGAGGCTAC CTGAGTGTG TACGAGGGGCC ACTGGCTCT CAGGAAGAGT TTTGTGTGA 720  
 TGGTTGAGCC CTAGTAGGGA CCAAGTGACA TCACATCCCT TCAAGATGCC TGAAGATCA 780  
 GCGAGTCTCT CTTCCTGCA GAGCTTTGGC CATTACACC TGACCTCTTG CTGCGAGCTA 840  
 ATAAAGAAATG CCAAGTGAGC AGTCTGGGCA CTGTCAAGGC AGGGAAGGGC CCAATGACTTT 900  
 TTGCGCTGC CCTCAAGCTG TTGCGCTGCC TCCAAACCC CATTAGTCTA GCTTGTGAGC 960  
 TGTATGCA AGTGTGTTT CTGGCTAGT CTGTTTCTTA AAGCAGAGAC TATTCCTCTT 1020  
 CTTCCGAGG AATATGTTT TCTCTTTGTC TTAATGATC TGTATGGGGA GAAATGGGGA 1080  
 ATGTCTACA CATGAGATGT TATATCTTG CGATGTACAG AATCAAGAGG TGGTTTGA 1140  
 GCATCATAA CAGGCTGACT GGCAGGAATG AAAAAAATA AAAAAAATA

Seq ID NO: 128 Protein sequence:  
 Protein Accession #: NP\_03247.1

1 11 21 31 41 51  
 MPSPRPAPS VVLLRLRLAL LRPPLGLEAC SCAPAHPPQH ICHISALVIRA KISSEKVVPA 60  
 SAPDADTERK LRYKLVYDQY QYITFBS LQVWLEBAS QKQYLLTG 120  
 LSDKGFPHL QNYIEPWEDL SLVQRSLNH HYHLNCGCQ/ TICTYVRCIT SAPNECLWTD 180  
 WLERKLYGY QAQHYVCMKH VDGTCISWYRG HLPRLKEFVD IVQP

Seq ID NO: 129 DNA sequence  
 Nucleic Acid Accession #: NM\_007207.2  
 Coding sequence: 143-1591 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CCACGCGTCC CCAATGAAGC CGAGTGAATG GGGGCTGAAT GTGCGAGTCC ATAGCTGAAG 60  
 AGCAGGCCA GAGATGGAG GAATGACACT ATTATGAACA GTGTCTGTGAG TTCTTCTTGA 120  
 ATTGCGAGTT TTACGCTCTCC TCATGCTCTC GTCTCTCTTA GACGACAGGG TAGTAGTGCC 180  
 ACTATCTAGG CCGGTCCGAC CTCAGGATCT CAACCTTTGT TTAGACTCTA GTTACTCTGG 240  
 CATTGCCAAC CCAGGCGATA ACAGCCACCC TCTGTCTATC GCCACGACCG TTGTCTCCCT 300  
 CAAGGCTGGC AATCTGACCT ATATGCGCTC ATCCAGCGGC TGTGCGCGCT CGTGAATTGT 360  
 TGGATGCAAC ACTGCCAGCT GCTGCACTGT GGCACACTAC GACAAGGACA ATCAGGCCA 420  
 AACCCAGCC ATTGCGCTG GGCACACAC CACTGTCCAT GGAAGCTCTA CCACTTGCCC 480  
 TGCTAACCA AGTGTCAACA ATAAATGAGA TACAGGCTCT CTAAGTCCAT CAGTGGGGGT 540  
 GGGCAGGCCCT GTGTCAAGGA CCCCCAAGCA GCTAGCCAGC ATCAAAATTA TCTACCCCAA 600  
 TGACTTGGCA AAGAAGATGA CCAATGCGAG CAAGAGATC ACCTGCCAGTC AGGGCCCTGT 660  
 CATCATTCAC TGCAGGCCCT TCATGGAGTA CACAAGAGAT CACATCCAAAG GAACTCTCCA 720  
 CATTAAGTGT GCGGCAAGA TCAGGCGCGG GAGACTGAGC CAGGCGAAGA TCACTGTCTC 780  
 AGACTTGATT TCTGTAGGAG AAGCGAAGCA CTTCTCAAG AGGAICTTT TCTTGGCAAA 840  
 TATAGTTTAT GATGAGAATA CCAATGAACC AAGCGGAGTG ATGCGCTGCC AGCCACTTCA 900  
 CATAGTCTCT GAGTCCCTGA AGAGAGAAGG CAAAGAACCT CTGGTGTTTA AAGGTGGAGT 960  
 TAGTAGTTTT AAGCAGAACC ATGAAGAACT CTTTGACAC AC TCCCTCCAGC TCCAGAGAGT 1020  
 CGGGAGAGTG GCGGCGCGCG CATCCGCGCG CTGAGCTCTG CTACTCTGAG CCAATCCGAC 1080  
 CACCGCTGAC ATGCGAAGCC CTGAGTCAAC CCGCACTCTG CCCTCTCTGT TCTTGGCAAA 1140  
 TGAGCGAGAT GCTCAGGACC TGGACACCAT CGAGCGGCTG AACATCGGT ACCTCATCAA 1200  
 GTCACCACT CATCTTCCC TCTACCACTA TGAGAAAGGC CTGTTCAACT ACAAGCGGCT 1260  
 GCGAGCCACT GACAGCAACA AGCAGAACTC GGGCGAGTAC TTGAGAGAG CTTTTGAGTT 1320  
 CATGAGGAA GCTCACAAGT GTGGGAAGG GCTTCTATC CACTGCGCAG CTGGGCTGTC 1380  
 CGCTGCCCG AGCATGTGA TCGTAACTT GATGAAGCAC ACTGGGATGA CCAATGACTGA 1440  
 TCTTTATAA TTTGTCAAAG GCAAAAGCAC AATTATCTCC CCAAACTCTA ACTTCATGG 1500  
 CGAGTTGCTA GAGTGTGAGG AAGACCTTAA CAACGCTGTG ACACCGAGAA TCTTTACAC 1560  
 AAAGCTGAGT GCGCTGAGGA CGGTTGTGTG ACAATGGTCT GGATGGAAAG GATGTCTGCT 1620  
 CTCCATTAGG AGACAATGAG GAAGAGAGAT GGATTCTGGT TTTTTCCTTT TCTTTTCTTT 1680  
 TTTGATGG GATTAAGTT TTGTAATGGA AACAACTGT GTTAAAGAT TATTTTATA 1740  
 CAGGTGTAAG AAGACTATAC TTTTGTAGCC ATTGAGATC ACCTTCCACA AACTGGCCAA 1800  
 ATAAAGGAGG TAAAGAAAGT AATTTTTTT AAGGCCAACC ATTAATAAT TAATACAACT 1860  
 TGGTTCTCC CCCTTTCTC TTTAAAGCTA NTTTGTAAAA GTTTATGAG

Seq ID NO: 130 Protein sequence:  
 Protein Accession #: NP\_009138.1

1 11 21 31 41 51  
 MPFSFLIDRRV VVALSPVVP QDLNCLDSS YLGSANFQSN SHEPPVIATT VSLKAAALTY 60  
 MPSSKQARS LNCQSSSAC CTVAATYKDN QAQTQAAAG TTTATGIST TCFANQAYVN 120  
 NENTUSLSPS SGVGSVSGT FKQLASIKI YPNDLAKMT KCSKSHLPSG GFVIDCRFP 180

MEYNKSHIQ AVHNCADKI SRRRLQGGKI TVLDLISCRE GKDSFKRIFS KSHVYDENT 240  
 NEPSKVMPSQ PLHIVLESKL REQKEPLVLK GGLSSFKQNH ENLNCDSLQL QCEKREVGGGA 300  
 SAASLLPQP IPTTDPDENA ELTPILPFLF LGNEBQDAQDL DTMQRLNGY VINVTIHLPL 360  
 VHYEKGLPNY KRLPATDSNK QNLRLQYFEEA PEPIEBAHQ GKGLLIHCQA GVSRSAITVI 420  
 AYLKHKTRMT MTDAYKFKVG KRPIISPILN FMQQLLEFE DLNNGVTPRI LTPKMGVET 480  
 VV

Seq ID NO: 131 DNA sequence  
 Nucleic Acid Accession #: NM\_005409.3  
 Coding sequence: 94-378 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 TTCTTTTCAT GTTCAGCATT TCTACTCCTT CCAAGAAGAG CAGCAAAAGCT GAAGTAGCAG 60  
 CAACAGCACCC AGCAGCAACA GCAAAAAACA AACAATAGTG TGAAGGGCAT GGCATATAGCC 120  
 TTGGCTGTGA TATGTGTGTC TACAGTTGTT CAAGGCTTCC CCAATGTTC AAGAGGACACC 180  
 TGCTTTTGA TAGCGGCTGG GGTAAAGACA GTGAAGTGG CAGATATTGA GAAAGCTTCC 240  
 ATAAATGAC CCAAGTAAACA CTGTGACAAA ATAGAAGTGA TTATTACCT GAAAGAAAT 300  
 AAAGGACAAC GATGGCTAAA TCCCAATCC AAGCAAGCA GGCTTATAAT CAAAAAGTT 360  
 GAAAGAAAGA ATTTTAAAA ATATCAAAAC ATATGAAGTC CTGAAAAAG GCATCTGAAA 420  
 AACCTAGAAC AAGTTTAACT GTGACTACTG AAATGACAAG AATTCTACAG TAGGAAACTG 480  
 AGACTTTTCT ATGGTTTGT GACTTTAACT TTGTGTACAG TTTGTGAG GATGAAAGGT 540  
 GGGTGAAGG ACCAAAAACA GAAATACAGT CTGTGTAAT GAATGACAAT CAGAATTCCA 600  
 CTGCCCAAG GAGTCCAGCA ATTAATATGA TTCTAGGAA AAGTCACTT AAGAAGAGCT 660  
 GTTACCATC GGAGTTTACA AAGTGCTTTC ACGTTCTAC TTGTGTATT ATACATTCAT 720  
 GCATTCTAG CTAAGAGAAC CTCTAGATT TGATGTACT AACATACTG TTGTGACTAT 780  
 GAGAACAATT GTTCTCTAG AAGTATCTG TCTGTATTGA TCTTATGCT ATATTACTAT 840  
 CTGTGTATC AAGTGAACA TTGACATT TACTGAGTC AAGGCTTAT AAGTGAAGA 900  
 CACTATATG TGTAAGAACA TTCTCAAAAC ATTTTTCAT GCAATATCAC ACTCTTTCC 960  
 CCAAAATCA TGTAGACACAT CAATATGTAG GGAACATCT TTATGCATCA TTGTGTGTT 1020  
 TTTATAACCA ATTCATTAAT TGTATTTCA AAAATGACT ATGAAAAAAA TTATACCTTA 1080  
 TGGGATAGT GCACACAGTGC AATATATTTCA TAACCAAAAT AGCAGCACCG GTCTTAATT 1140  
 GATGTTTTC AACTTTTAT CATATATTTG TTTGAACA ATTAGGATAT GTGTGTAC 1200  
 TGTACTTTT GTTTTGA TCT GTTTGTATA ATGATAGCAA TATCTTGAC ACATTGTGAA 1260  
 TACAAAATGT TTTGTCTAC CAAAGAAAAA TGTGAAAAA TAAGCAAATG TATACCTAGC 1320  
 AATCACTTT ACITTTTGA ATTCTGTCT TTAGAAAAAT ACATAATCTA ATCAATTCT 1380  
 TTGTCATGC CTATATACTG TAAATTTAG GTATACTCAA GACTAGTTA AAGAATCAAA 1440  
 GTCATTTTT TCTGATAAT ACTAACCAAA CTTTCTTTT TTAAAAAAAA AAA

Seq ID NO: 132 Protein sequence:  
 Protein Accession #: NP\_005400.1

1 11 21 31 41 51  
 MSVKGMAIAL AVILCATVVO GPFMKRGRCL LGPGVKAV KVADIEKASI MYPNNCDKI 60  
 EVITLKENK GQRCLNPKSK QARLIKKVE RKNF

Seq ID NO: 133 DNA sequence  
 Nucleic Acid Accession #: NM\_012342  
 Coding sequence: 373-1153 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 CTGGCCCGGG CCGGAGCTGC GCGGTATAC: CTTCGTGCT GTGGAGCCC TACTCTCTTC 60  
 GCTGGAAGCG GCGGCTATG GCGGAGGACC GACTCCGAC CCGGGGCTTAC 120  
 CTGTGCTGCC TAGACTCGAG CGGGGCAACG GAGCCAGTGG CCGGCGACGG GGGACCGGGA 180  
 AACCTTTCTG GCTCTCGGA GAGCCCTGTA GCGCGCTCC ATGCTCCGC AGCGGCCGGA 240  
 AACCCAGGCC CCGCGCTGAC GGAGCCCGCG CCTCCGGCCA GGGCCCATGC CCTCGCGCT 300  
 CCGGGGGTGG TAGCTGTGCG CGAGCCGGGG CTTCGGAAG CCGCGGGGCG CCGCGGGGCG 360  
 TGCGGGGCGT CAATGTGCT CCACTCAGCG TACATCTGA TCTGGCTGCA CTGGAGGCTC 420  
 TGCCCATCG CCGTGTGCT GACCAAGGT GAAATTCGAT GCTACTGTGA TGCTGCCACC 480  
 TGTTAGGACA CTGGTTATA T GTGTAAAT GAGTCAAGCG CCGTCTCTCT TGAATCTTT 540  
 GATCCTCAGA ACTCAAACT CCGACTCACC CATGGCTGCC TGGACTCTT TCGAAGCAGC 600  
 ACACAGCATCT GCCAAGCCAA ACAGGGCCGGA AACCACTGT GCACACACAT ACCGACATTG 660  
 GAATGCTGTC ATGAAGACAT GTGCAATTG AGAGGGCTGC ACGATTTCT CTCTCCGCC 720  
 AGGGGTGAG CCGTAGGCA AGGAACAGG TATCAGCAT ATGGTAGCAG AAACCTTATC 780  
 ACCAAGGTGC AGTAGCTGAC TTCTTCAAA GAGTGTGGT TCGGGGACG GGTTCATTGC 840  
 GTGCCCATTC CTGGAGGCTT GATTTTAGTG TTGCTTATTA TGTGGCCCT GAGGATGCTT 900  
 CGAAGTGAAA ATAAAGAGCT GCAGGATCAG CCGCAACAGA TGCTCTCCG TTGACTAC 960  
 AGCTTTCAG GCACACATTC CAAAAGGCG CAGGTTGCAA AGTTAGACT GGAATGCATG 1020  
 GTGCCGCTCA GTGGCGAGCA GAACCTGCT GTGACTGTG ATAAAATGAC ACAAGGAGAC 1080  
 CTCACCAAG ATAAAGATCT CTGCTTGT TTCTAGTGGCA TGTACAGTGG GCACGGGAAG 1140  
 CTGGAAATTC TATGACGGAG TCTTATCTGA ACTACACTTA CTGAACAGCT TGAAGGCTT 1200  
 GTGATTCTG CTGGACAGGA GCACTTTATC TGAAGACAAA CTCATTTAAT CATCTTTGAG 1260  
 AGACAAAATG ACCTCTGCAA ACAGAAATCT GGAATTTCT TCTGAAGGAT TATTTGACA 1320  
 GACTTAAATA CAGTAAATG TGTATTGCT TTTTAAAT ATAAAAGCA AGGAGAGAC 1380  
 TTGTACACA CTGTACAGG GTTATTGTC ATCCAAAGGA CTGGGAATG AGTACTTAAA 1440

TAAACAAAAA TGTGCCATAT GTAAGCTTCT ACATCTGTAT TTAITGTAAA GATTITAAAAA 1500  
AAATATATAT ATTTTGTCTG A

Seq ID NO: 134 Protein sequence:  
Protein Accession #: NP\_036474.1

1 11 21 31 41 51  
MDRHSSVIRI WLQLELCAMA VLLTKGEIRC YCDAAHCVAT GYMCKSELSEA CFSRLDPQN 60  
10 NSNPLTHIGLC DLSASTTIDIC QAKQARNHNS TTITFLECH EDMCNYRGLH DVLSPRGEA 120  
SQQGNRYQHD GSRNLITKVF QLTSSKELWF RAAVIAVPIA GLGILLVLLM LALRMLRSEN 180  
20 KQLDQDRQQM LSRLLYHSFG HHSSKGGQVAK LDLECMVPVS GHENCLCTD KMRQADLSND 240  
KILSLVHWGM YSGHGKLEFV

Seq ID NO: 135 DNA sequence  
Nucleic Acid Accession #: NM\_001627.1  
Coding sequence: 64-1815 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
CGGACGACG CCCCCTCTCTG CGGCGTGAC TCGTCAGTG GCCACCAAG AAGGAGGAGG 60  
AATATGGAAT CCAAGGGGGC CAGTTCCTGC CGTCTGCTCT TCTGCTCTCT GATCTCCGCC 120  
AOCGTCTICA GGCACCGCTT TGGATGTTAT ACTGTAAATC TACGATATGG AGATACCACT 180  
ATCATCTCTT GCGCATCTGA GATCTCATGT TTGCGCAATG GAATATATGA 240  
AAGCCCGAGT GCTCCCGGCT ATTTATTGGC TTAGCATCTC GTACAAAGAA AAGTGTGCGAG 300  
TAGCAGCATG TACCAGATA CAAAGACAGA TTGAACCTCT CAGAAAACCT CACTTTGTCT 360  
ACTAGTAATG CAAGGACTCAG TGATGAAAG AGATTITGTG GCATGCTAGT AACTGAGGAC 420  
AAGCTGTTTG AGGCACCTAC AATAGTCAAG GTGTGCAAGC AACCATCTAA ACCTGAAATC 480  
GTAAAGCAAG CACTGTTTCT CGAAACAGAG CAGCTAAAAA AGTTGGGTGA CTGATTTICA 540  
GAAGACATG ATCCAGATGG CAATTATACA TGGTACAGA ATGGAAAAAG GTCATACACC 600  
CTTGAAGGAG CCGTGGTCAAT AATTTTAA AAGGAAATGG ACCCAGTGAC TCAGCTCTAT 660  
AACATGACTT CCACCTGGG ATACAAGACA ACCAAGGCTG ACATACAAAT GCAATTCACC 720  
TGCCTGGTGA CATATTATGG ACCATCTGCG CAGAAAACAA TTCACTCTGA ACAGGACGTA 780  
TTTGATATTT ACTATCTTAC AGAGCAGGTG ACAATACAAG TCTGCCACC AAAAATGTCT 840  
ATCAAGAAAG GGGATACAT CACTCTTAA TGCCTAGGA ATGGCAACC TCCCGAGAG 900  
GAATTTTGT TTACTTACC AGGACAGGCC GAAGGAATAA GAAGCTCAA TACTTACAA 960  
CTGATGGATG TGAGGCGCAA TGCAACAGGA GACTACAAGT GTTCCGTGAT AGACAAAAA 1020  
AGCATGATG CTTCACACAG CATCACAGTT CACTATTGCG ATTGTGCTT AAACCCAACT 1080  
GGGAAAGTGA CTAGACAGAT TGGTGAATGC CTACCGGTGT TATGACATAAT ATCTGTAGC 1140  
AGGATACAA CIGTGATATG TGGGAAAGAT AACATCAGGC TTGATATGAG CCGGTGATT 1200  
TTTAGCTTTC ATTATACGA TCGTGGAAAC TATGTCTGCG AAATGCTCT GCAAGGAGTT 1260  
GAAGGACTAA AGAAAAGAGA GTCAATTGACT CTCATTGTAG AAGGCAAAAC TCAATATAA 1320  
ATGACAAAGA AAATGATCC CAGTGGACTA TCTAAACAAA TAACTTGCCA TGTGGAAAGT 1380  
TTTCCAAAGC CAGCCATTCA GTGGACAATT ACTGGCAGTG GAAGCGTCAT AAACCAACAA 1440  
GAGGAATCTC CTATATTAA TGGCAGGTAT TATGATAAAA TTATCATTTG CCTGGAAGAG 1500  
AATGTACAT TAACTGTCAC AGCAAGAAAC CAACTGGAGA GAACAGTAA CTCCTGGAAT 1560  
GTCTCTGCTA TAAGTATTCC AGAACAGAT GAGGCGAGCG AGTAAAGTGA TGAACAACGA 1620  
GAAAAGGTGA ATGACAGGCG AAAACTAAT GTGGGAATGG TTGTTGTCT CTCTCTGCT 1680  
GCCCTGTGTG CTGTGTCTGT CTACTGGCTG TACATGAAGA AGTCAAAAGC TGCATCAAAA 1740  
CTGTATAAACA AGGACCTCGG TAA TATGGAA GAAACAAAAA AGTTAGAAAG AACAATCATC 1800  
AAAGCTGAAG CCTAAGAGG AAACCTGTCT AGTTTCCGAG AGATAAAAT CATATAGACC 1860  
AATTGAGCA TGAAGCTGTA TTGTATTAA GACATACAAA AAGACATGGA CAGCAATICA 1920  
TGTTTCAGT ATTAAGCAGT TCAATCTACC AAGCTGTAC AGGTTTCAG AGAATATCT 1980  
CAGGTAAAAA AATGAAGATT TAATTACAAA CATAAAGAAC AAGTTTGGC AGCCATGATA 2040  
ATAGGTCATA TGTGTGTTT GGTTCAAATT TTTTCCGTGA AATGTCTGCA CTGAGGATTI 2100  
CTTTTGGT TTGCTTTTAT GTAAATTTT TACGTAGCTA TTTTATACA CTGTAAGCTT 2160  
TGTCTGGGA GTTCTCTTA ATCTATGTA TAATGTATG TTTTATTCT AATTGTTTAT 2220  
ATGGATAATC TGAGCAGGTA CATTCTGAT TCTGATTGCT ATCAGCAA TGCCAAACT 2280  
TCTCATAAGC ACCTAAAAAC CAAAGGTGGC AGCTTGTGAA GATTGGGAC ACTCATATTG 2340  
CCCTAATTA AACTGTGAT TTTATCACA AGGAGGGGA GCGCCAGAGT CAGACTGATA 2400  
GAGCAGCATG GAGCGGACTT TTGATATGC CACCGAGCAA CTTCAAGAAA TAAATTCAG 2460  
ATGATATAG CAGACATAC ATAAATGTAT TCCCAACTG ACAATTTTAC CTATTCTGAA 2520  
AAGGACATA AACAGAAAT

Seq ID NO: 136 Protein sequence:  
Protein Accession #: NP\_001618.1

1 11 21 31 41 51  
MESKGASSCR LFLCLLSAT VFRPGLOWYT VNSAYGDTTI IPRCLDVPQN LMFQWKYKE 60  
70 PDOSVFPIAF RSTKSKSVQY DDVPEYKDLRL NLSNRYTLI SNARISDEKR FVCMILVTEN 120  
VHEAFVTKVY FQKSKSFVSI SKLLELETEG KGLLDGISE DSYPDNIVY YRNSKVLFE 180  
EGAVVHFKK EIVFVQQLY MTSLRYEYCT KADQMPFTC SVTVYPSQK XTHISSQAV 240  
DLYVFTEYQT IQVLPKNAI KEGDNITLKC LIGNNPPPE FLFYLQQFE GRHSNNYTL 300  
75 MDVRRNATGD YKCLSIDKKS MIASTAITVH YLDLSNFGS EVTRQIGDAL PVSTISASR 360  
NATVYVMKDN IRLRSFSPFS SLHYQDAGNY VCETALQBE GLKRESLRL IVEGPKQDQM 420  
TKKTPDSGLS KTIICHVBF PKPAIQWTH GSGSVINQTE ESPYNGRYV SKLISPEN 480

VTLTCTAENQ LERTVNSLV SAISIPSHDE ADHSIDENKE KYNDQAKLIV GIVVGLLLAA 540  
 LVAGVVYWL YMKKSKTASKE VNKDLGNMEE NCKLENNHK TEA

Seq ID NO: 137 DNA sequence  
 Nucleic Acid Accession #: XM\_050359  
 Coding sequence: 1-119 (underlined sequences correspond to start and stop codons)

1 11 21 31 41 51  
 | | | | |  
 ATGAACCGCA GCCACCGGCA CGGGGGGGGC AGCGGCTGCC TGGGCACATAT GGAGGTGAAG 60  
 AGCAAGTTTG GAGCTGAATT TCGTGGTTT TCGCTGGAAA GATCAAAACC TGGAAATTT 120  
 GAGCAATTTT ATGGAATCTT ACAACA TGTT CATAAGATCC CCAATGTTGA CTTTGTGTA 180  
 GGCTATGCAG ACATCCATGG AGACTTACTA CCTATAAATA ATGATGATAA TTATCAAAA 240  
 GCTGTTTCA CGGCCAATCC ACTGCTTAGG ATATTATAC AAAAGAAGGA AGAAGCAGAC 300  
 TACAGTGCCT TTGATACAGA CACGCTAATA AAGAAGAAGA ATGTTTAAAC CAACGATTG 360  
 CTTCTGACA ACCATAGAAA AAAGCCACAT ATATTCATTA GTATGCCCA AGACTTTAGA 420  
 CTTGTGCTT CTATTATAGA GTTGAATAT TTCCAGAAA CGCATCTAG GGTACGCTT 480  
 TACAAATAGC GCACGGAGAA ACCCTAGGA TTCTACATCC GGTATGGCTC CAGTGTACAG 540  
 GTAACACCAC ATGGCTTAGA AAGGTTTCCA GGGATCTTTA TATCCAGGCT TGTCCACGA 600  
 GUTCTGGCTC AAAGTACAGG ACTATTAGCT GTTAATGATG AAGTTTGA AGTTAAATGGC 660  
 ATAGAAATTT CAGGGAAGAG CTTGATCAA GTAACAGACA TGATGATTGC AAATAGCCCT 720  
 AACCTCATCA TAACAGTAGG ACCGCGAACC CAGAGGAATA ATGTTGTGAG GAACATGCG 780  
 ACTCTGCA GTTCGGTCA GTCTACTGAT AACAGCCTTC TTGCTACCC ACACGAGATT 840  
 GAACCAAGCT TTGAGCCAGA GGTATGAAGC AGCGAAGAAG ATGACATTAT CATTGAAGAC 900  
 AATGGAATGC CACAGCAGAT TCCAAGAGCT GTTCTAATA CTGAGAGCCT GGAGTCATTA 960  
 ACACAGATAG AGCTAAGCTT TGAGTCTGGA CAGAAATGCT TTAATCCCTC TAATGAAGTG 1020  
 AGCTTAGCAG CCAATAGCAAG CAGCTCAAAAC ACAGAAATTT AAACACATGC TCCAGATCAA 1080  
 AAACCTTTAG AAGAAGATGG AACAAATCAT ACATTATGA

Seq ID NO: 138 Protein sequence:  
 Protein Accession #: XP\_030359

1 11 21 31 41 51  
 | | | | |  
 MNRSRHRHGA SOCLGTMIEVK SKFGAHRFRF SLERSKPGKF EEFYGLLQHV HKIPNVDVLV 60  
 GYADHHDLL PNNDNDNYHK AVSTANPLLR IFIQKKBED YSAFGDTLL KKKNVLTNVL 120  
 RPDNHRKKPH IVISMPQDFR PVSSMDYDI LPETHRRVRL YKYGTGKPLG FYIRDGSSVR 180  
 VTTHLEKVP GFISLVPFG GLAQSTGLA VNDVLEVNG IEVSGKSLDQ VITDMNANR 240  
 NLITVRAN QNRYVYRNSK TDSGSGQSD NSLLGYPOQ EPSPFEDD SEEDDIED 300  
 NGVFPQPKA VNTESLES L TQELSFESG QNGFIPSNV SLAAISSN TEFETHAFDQ 360  
 KLLBEDGTL TL

It is understood that the examples described above in no way serve to limit the true scope of this invention, but rather are presented for illustrative purposes. All  
5 publications, sequences of accession numbers, and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

WHAT IS CLAIMED IS:

- 1                   1.       A method of detecting a breast cancer-associated transcript in a cell  
2       from a patient, the method comprising contacting a biological sample from the patient with a  
3       polynucleotide that selectively hybridizes to a sequence at least 80% identical to a sequence  
4       as shown in Tables 1-25.
- 1                   2.       The method of claim 1, wherein the biological sample comprises  
2       isolated nucleic acids.
- 1                   3.       The method of claim 2, wherein the nucleic acids are mRNA.
- 1                   4.       The method of claim 2, further comprising the step of amplifying  
2       nucleic acids before the step of contacting the biological sample with the polynucleotide.
- 1                   5.       The method of claim 1, wherein the polynucleotide comprises a  
2       sequence as shown in Tables 1-25.
- 1                   6.       The method of claim 1, wherein the polynucleotide is immobilized on  
2       a solid surface.
- 1                   7.       The method of claim 1, wherein the patient is undergoing a therapeutic  
2       regimen to treat breast cancer.
- 1                   8.       The method of claim 1, wherein the patient is suspected of having  
2       breast cancer.
- 1                   9.       An isolated nucleic acid molecule consisting of a polynucleotide  
2       sequence as shown in Tables 1-25.
- 1                   10.      The nucleic acid molecule of claim 9, which is labeled.
- 1                   11.      An expression vector comprising the nucleic acid of claim 9.
- 1                   12.      A host cell comprising the expression vector of claim 11.

- 1           13.    An isolated polypeptide which is encoded by a nucleic acid molecule  
2   having polynucleotide sequence as shown in Tables 1-25.
- 1           14.    An antibody that specifically binds a polypeptide of claim 13.
- 1           15.    The antibody of claim 14, further conjugated to an effector component.
- 1           16.    The antibody of claim 15, wherein the effector component is a  
2   fluorescent label.
- 1           17.    The antibody of claim 15, wherein the effector component is a  
2   radioisotope or a cytotoxic chemical.
- 1           18.    The antibody of claim 15, which is an antibody fragment.
- 1           19.    The antibody of claim 15, which is a humanized antibody
- 1           20.    A method of detecting a breast cancer cell in a biological sample from  
2   a patient, the method comprising contacting the biological sample with an antibody of claim  
3   14.
- 1           21.    The method of claim 20, wherein the antibody is further conjugated to  
2   an effector component.
- 1           22.    The method of claim 21, wherein the effector component is a  
2   fluorescent label.
- 1           23.    A method for identifying a compound that modulates a breast cancer-  
2   associated polypeptide, the method comprising the steps of:  
3       (i) contacting the compound with a breast cancer-associated polypeptide, the  
4   polypeptide encoded by a polynucleotide that selectively hybridizes to a sequence at least  
5   80% identical to a sequence as shown in Tables 1-25; and  
6       (ii) determining the functional effect of the compound upon the polypeptide.
- 1           24.    A drug screening assay comprising the steps of



- 2 (i) administering a test compound to a mammal having breast cancer or a cell  
3 isolated therefrom;
- 4 (ii) comparing the level of gene expression of a polynucleotide that selectively  
5 hybridizes to a sequence at least 80% identical to a sequence as shown in Tables 1-25 in a  
6 treated cell or mammal with the level of gene expression of the polynucleotide in a control  
7 cell or mammal, wherein a test compound that modulates the level of expression of the  
8 polynucleotide is a candidate for the treatment of breast cancer.